ELSEVIER

Information Sciences 132 (2001) 23–51

INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL

www.elsevier.com/locate/ins

Optimization design of biorthogonal filter banks for image compression

Yi Shang [a,*], Longzhuang Li [a], Benjamin Wah [b]

[a] *Department of Computer Engineering and Computer Science, University of Missouri-Columbia, Columbia, MO 65211, USA*

[b] *Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*

Abstract

In this paper, we present a new approach for designing filter banks for image compression. This approach has two major components: optimization and generalization. In the optimization phase, we formulate the design problem as a nonlinear optimization problem whose objective consists of both the performance metrics of the image coder, such as the peak signal-to-noise ratio (PSNR), and those of individual filters. Filter banks are optimized in the optimization phase based on a set of training images. In the generalization phase, the filter bank that can be generalized to other images is selected from the candidates obtained in the optimization phase to be the final result. The filter bank selected should perform well not only on the training examples used in the design process but also on test cases not seen. In contrast to existing methods that design filter banks independently from the other operations in an image compression algorithm, our approach allows us to find filter banks that work best in a specific image compression algorithm for a certain class of images. In system prototype development, we adopt the agent-based approach to achieve better modularity, portability, and scalability. Agents in the multi-agent system are specialized in performing problem formulation, image compression, optimization, and generalization. In the experiments, we show that on a set of benchmark images our approach has found filter banks that perform better than the existing filter banks in different image compression algorithms and at different compression ratios. © 2001 Published by Elsevier Science Inc.

* Corresponding author. Tel.: +1-573-882-3371; fax: +1-573-882-8318.
E-mail address: shangy@missouri.edu (Y. Shang).

0020-0255/01/$ - see front matter © 2001 Published by Elsevier Science Inc.
PII: S0020-0255(01)00067-6

Keywords: Nonlinear optimization; Generalization; Image compression; Multi-agent systems

1. Introduction

The design of digital filter banks is important because filter banks have been used in many applications, including modems, data transmission, speech and audio coding, and image and video coding [16,46,55]. In this paper, we present a new approach for designing filter banks that is based on optimization and generalization. We apply the approach to the design of two-band biorthogonal filter banks for subband image coding and show promising results. Our design method is general enough to be applicable to the design of other types of filter banks, such as multi-rate and multi-band filter banks.

Recently, transform coding has become the de facto standard for image and video compression. It is based on the principle that a transformed image has better energy compaction and decorrelation properties that make the image easier to compress. Generally speaking, the transformations can be either linear or nonlinear. Linear transformations such as the *discrete cosine transform* (DCT) and the *discrete wavelet transform* (DWT) are used more often in image compression because of their invertible property. These transformations are implemented efficiently in subband image coders in the form of digital filter banks.

Subband image coding generally consists of three phases: subband transform, bit allocation and quantization, and lossless coding such as entropy coding [2,48]. In subband transform, a filter bank is applied to split the input signal into the low-frequency reference signal and the high-frequency detail signal. The transformation can be applied to the reference and detail signals recursively. Usually, the number of this splitting ranges from 3 to 5, depending on the target compression ratio, the size of the original image, and the length of filters. After the recursive subband transform, a set of transform coefficients are generated that represent subimages of various sizes. All information in the original image is preserved in subband coefficients and no compression has been done so far. A perfect-reconstruction filter bank does not introduce any errors by itself.

In the quantization and bit allocation phase, compression is achieved by compacting and throwing away some of the less important information. Subband transform splits the original image into subimages corresponding to information of difference importance to the human visual system. Important information needs to be preserved while less important information may be discarded in situations such as low bandwidth communications. Based on the

information contained in subimages, quantization and bit allocation methods assign various number of bits to the quantized subband coefficients in subimages so that essential information is preserved in a limited number of bits.

Finally, subband coefficients are coded efficiently to achieve a high compression ratio. Traditional coding methods include entropy coding and arithmetic coding. In recent years, a number of wavelet coding methods have been proposed to improve the compression ratio significantly by exploiting the relationship of the coefficients across subbands [11,32]. Examples include embedded zerotree method [39], set partitioning in hierarchical tree (SPHIT) method [36], and significance-linked connected component analysis (SLCCA) method [7].

Biorthogonal filter banks have the best performance among various types of filter banks in image compression. Perfect-reconstruction biorthogonal filter banks are widely used because of the nice property that they do not introduce any errors by themselves, which is necessary for lossless image coding. However, in lossy image coding situations such as in very low bit rate compression, some information has to be discarded in the compression process, either in the transformation represented by filter banks or somewhere else. In this case, perfect reconstruction (PR) is unnecessary and small reconstruction errors introduced by filter banks is tolerable.

Many methods have been developed to design filter banks for subband image coding. These methods can be classified into two categories. One approach designs a filter bank independently, without considering a specific image compression algorithm and a particular set of images. By carefully choosing design criteria, this approach can find good solutions such as the Antonini's spline (9,7) wavelet filter [2,4,48]. The drawback of this approach is that solutions obtained may not perform the best for a specific class of images. The other approach optimizes compression performance by designing filter banks in the context of a specific image coding algorithm based on training images [5]. Hence, the filter banks designed are image coder dependent and image dependent, and they usually work better for the specific set of training images. The drawback of this approach is that the results may not be generalizable to images not in the training set.

In this paper, we focus on the design of optimal biorthogonal filter banks for image compression. We present a new design approach that consists of two major components: optimization and generalization. In the optimization phase, we formulate the design problem as a nonlinear optimization problem whose objective consists of both the performance metrics of the image coder, such as the *peak signal-to-noise ratio* (PSNR), and those of individual filter such as the stopband and passband energies of a low-pass filter. Filter coefficients are searched to optimize the performance metrics on a fixed set of training images by using nonlinear optimization methods such as simulated annealing. One or a few optimal/near-optimal coefficient sets are obtained

based on individual training image. In the generalization phase, one coefficient set is selected from the results generated in the optimization phase as the final result. The set selected is generalizable, meaning that it performs well not only on the training cases used in the optimization but also on test cases not seen. Our approach allows us to find optimal filter coefficients for a specific image compression algorithm on the target class of images.

In system prototype development, agent-based approach is adopted for achieving better modularity, portability, and scalability. The agents are specialized to carry out different roles in the multi-agent system. They are generalization agents, problem formulation agents, optimization agents, and image compression agents. The generalization agent performs problem decomposition in the beginning and integration and generalization of the final result in the end. The problem formulation agent builds the optimization model and supplies the objective and constraints to the optimization agent. The optimization agent searches for optimal filter coefficients by applying appropriate optimization methods. Finally, the image compression agent evaluates the image compression performance by running the image compression program on sample images. In the experimental results, we show that our approach found filter banks that perform better on a set of benchmark images than the existing filter banks in different image compression algorithms and at different compression ratios.

This paper is organized as follows. In Section 2, we introduce biorthogonal filter banks and various performance metrics. In Section 3, we present our new design approach and discuss related issues on optimization and generalization. In Section 4, we present a multi-agent framework for implementing the new design approach. In Section 5, we report our experimental results on benchmark images. Finally, in Section 6, we conclude the paper.

2. Biorthogonal filter banks for image compression

Fig. 1 shows a two-band filter bank that first decomposes an input signal $x(n)$ in time domain into two frequency subbands, and then combines them into a single output signal $\hat{x}(n)$. The filter bank consists of an analysis stage and a synthesis stage. In the analysis stage, the input signal is divided into two equal subbands using the analysis filters $H_0(z)$ and $H_1(z)$. According to the Nyquist Theorem, each subband signal is then down-sampled by 2 to form the outputs of the analysis stage. These signals can then be analyzed or processed in various ways depending on the application. For example, in a subband coder for audio or image compression, these signal are quantized, encoded and transmitted to the receiver. When there is no error occurred in the middle, the input signal to the synthesis stage is the same as the output of analysis stage. In the synthesis stage, each subband signal is up-sampled by 2 to form $f_0(n)$ and $f_1(n)$,

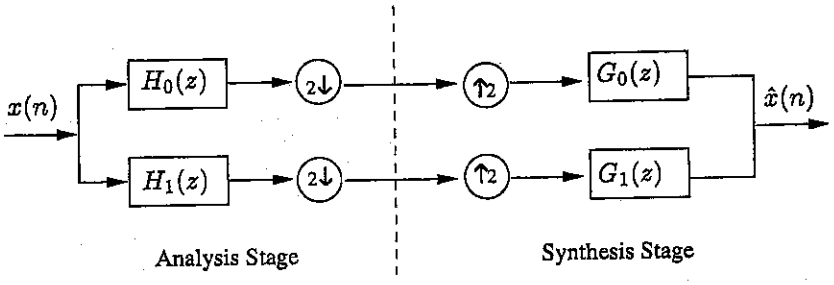

Fig. 1. A two-band filter bank that decomposes an input signal $x(n)$ into two subsignals in the analysis stage and later combines the signals into reconstructed signal $\hat{x}(n)$ in the synthesis stage.

processed by the synthesis filters $G_0(z)$ and $G_1(z)$, and summed at the output to form the reconstructed signal $\hat{x}(n)$ [1,30,45].

The Z-transform of the reconstructed signal, which is a function of $x(n)$ and the filters, is as follows:

$$\begin{aligned}\hat{X}(z) &= \frac{1}{2}([G_0(z)H_0(z) + G_1(z)H_1(z)]X(z) + [G_0(z)H_0(-z) \\ &\quad + G_1(z)H_1(-z)]X(-z)) \\ &= T(z)X(z) + S(z)X(-z). \end{aligned} \tag{1}$$

There are three types of undesirable distortions in a filter bank. *Aliasing distortions* include aliasing caused by subsampling and images caused by upsampling. The aliasing term $S(z)X(-z)$ in (1) represents the aliasing distortion. *Amplitude distortions* represent deviations of the magnitude of $T(z)$ in (1) from unity. *Phase distortions* represent deviations of the phase of $T(z)$ from the desired phase property, such as linear phase.

Extensive research has been conducted to remove undesirable distortions of filter banks. Aliasing distortions can be removed by selecting synthesis filters based on analysis filters. Infinite-impulse-response (IIR) all-pass filters can be used to eliminate magnitude distortions, whereas a linear-phase finite-impulse-response (FIR) filter removes phase distortions. PR of the original signal by a filter bank requires $S(z) = 0$ for all z, and $T(z) = cz^{-d}$, where c and d are constants. constant for scaling the magnitude of the output signal. Therefore, to perfectly reconstruct a signal, the transfer function is a pure delay with no aliasing, no amplitude change, and linear phase.

Criteria for the design of filter banks are application dependent. For instance, in subband image coding, it is desirable to make analysis filter $H_0(z)$ a low-pass filter and $H_1(z)$ a high-pass filter so that the frequency band can be divided into low and high frequencies. The stopband energy and passband energy of a filter can be used to measure the proximity to the ideal step filter.

Table 1
Possible design objectives of filter banks for image compression come from three sources: performance metrics of image compression, of the overall filter bank, and of a single filter

	Image compression	Overall filter bank	Single low-pass filter
Performance metrics	Signal-to-noise ratio Compression ratio	Amplitude distortion Aliasing distortion Phase distortion	Stopband ripple Passband ripple Transition bandwidth Stopband energy Passband flatness

Table 1 summarizes various performance metrics that can be used as filter-bank design objectives. The performance metrics for image compress include signal-to-noise ratio and compression ratio; The metrics for the overall filter bank include amplitude distortion, aliasing distortion, and phase distortion; And the metrics for a single low-pass filter include stopband ripple, passband ripple, transition bandwidth, stopband energy, and passband flatness [51,52].

2.1. Design criteria of a filter bank

PR and linear phase are two desirable properties of the filter banks used for image compression. PR avoids introducing errors by the filter bank, whereas linear phase prevents visible phase distortion around edges of an image.

In image coding, quadrature-mirror-filter (QMF) filter banks and biorthogonal filter banks are among the most widely used filter banks. In QMF filter banks, the low-pass analysis filter is the prototype filter from which the other three filters can be derived [9,10,18,22,25]. In principle, the two analysis filters are chosen to form a mirror pair in the frequency domain and the synthesize filters are decided accordingly to remove aliasing. The advantage of QMF filter banks is that only one filter needs to be designed. However, theoretically QMF filter banks with filter longer than two are not able to obtain Pr. In addition, QMF filter banks have to use much longer filters than those in biorthogonal filter banks in order to get comparable performance in image coding.

In biorthogonal filter banks, the two analysis filters form the basis of the design, and the two synthesis filter can be derived from them. PR and linear phase of filters is achieved when the analysis and synthesis filters are designed as follows [18,30,48].

First, choose the synthesis filters $G_0(z)$ and $G_1(z)$ based on the analysis filters to remove aliasing. For example, when the two synthesis filters are chosen as

$$G_0(z) = H_1(-z) \quad \text{and} \quad G_1(z) = -H_0(-z), \tag{2}$$

the aliasing term in Eq. (1) is 0 and the transfer function becomes

$$T(z) = (H_0(z)H_1(-z) - H_1(z)H_0(-z))/2. \tag{3}$$

Second, let the two analysis filters $H_0(z)$ and $H_1(z)$ satisfy the following conditions:

- The total length of them is a multiple of 4.
- They are FIR filters and are of even length or odd length at the same time. When they are of even length, filter $H_0(z)$ is symmetric and filter $H_1(z)$ antisymmetric. When they are of odd length, both of them are symmetric.
- To make the transfer function a pure delay, i.e. $T(z) = z^{-d}$ for a constant d, the coefficients of $H_0(z)$ and $H_1(z)$ need to satisfy a set of equations called the PR condition. When both filters $H_0(z)$ and $H_1(z)$ are symmetric and of odd length, the PR condition is [18]

$$\frac{1}{2}\theta\left(i - \frac{N_0 + N_1}{4}\right) = \sum_{k=1}^{2i}(-1)^{k-1}h_0(2i+1-k)h_1(k)$$
$$\text{for } i = 1, 2, \ldots, \frac{N_0 + N_1}{4}, \tag{4}$$

where $h_0(n), n = 1, \ldots, N_0$, are coefficients of $H_0(z)$ with length N_0; $h_1(n), n = 1, \ldots, N_1$, are coefficients of $H_1(z)$ with length N_1; and $\theta(x) = 1$ if $x = 0$, and 0 otherwise.

2.2. Design criteria of a single filter

Although there exists many biorthogonal filter banks with PR, Only a few biorthogonal filter banks perform well in image compression [2,35,48]. Criteria besides perfection reconstruction have been studied in order to identify good filter banks for image compression.

Based on the subband coding theory and the wavelet theory, metrics of individual filters in biorthogonal filter banks have been used in forming new criteria. A widely used criterion in subband coding theory is to make the two analysis filters approximate the ideal low-pass and high-pass filters, respectively. A pair of ideal filters can perfectly divide the frequency band into low and high frequencies. The amplitude response of an ideal low-pass filter is 1 in the frequency range from 0 to $\pi/2$, and is 0 in the range from $\pi/2$ to π. On the other hand, the amplitude response of an ideal high-pass filter is 0 in the frequency range from 0 to $\pi/2$, and is 1 in the range from $\pi/2$ to π. Since it is impossible to achieve these ideal situation using finite-length FIR filters, we can only approximate the ideal case.

The performance metrics for measuring the proximity to an ideal low-pass or high-pass filter include stopband ripple, passband ripple, transition bandwidth, stopband energy, and passband energy. Among these metrics, stopband

energy and passband energy are often used because they have closed-form formulations whereas the other three do not. The stopband energy and passband energy of a low-pass filter can be derived analytically as follows [51].

Assuming an odd-length symmetric low-pass filter $H_0(z)$ with coefficients $h_0(n), n = 1, \ldots, N_0$, the Fourier transform of $H_0(z)$ is

$$F_0(e^{jw}) = H_0(w)e^{-j\frac{N_0-1}{2}w}, \tag{5}$$

where

$$H_0(w) = h_0\left(\frac{N_0+1}{2}\right) + \sum_{n=1}^{(N_0-1)/2} 2h_0(n)\cos\left(\frac{N_0+1}{2} - n\right)w. \tag{6}$$

Then based on the definition, the stopband energy $E_s(h_0)$ with stopband cut-off frequency w_s is

$$\begin{aligned} E_s(h_0) = \int_{w_s}^{\pi} H_0^2(w)\,dw &= h_0^2((N_0+1)/2)(\pi - w_s) - 4h_0((N_0+1)/2) \\ &\times \sum_{n=1}^{(N_0-1)/2} h_0(n)\frac{\sin((N_0+1)/2 - n)w_s}{(N_0+1)/2 - n} \\ &+ 2\sum_{n=1}^{(N_0-1)/2} h_0(n)^2\left[(\pi - w_s) - \frac{\sin(N_0+1-2n)w_s}{N_0+1-2n}\right] \\ &- 2\sum_{n=1}^{(N_0-1)/2}\sum_{m=1,m\neq n}^{(N_0-1)/2} h_0(n)h_0(m) \\ &\times \left[\frac{\sin(n-m)w_s}{n-m} + \frac{\sin(N_0+1-n-m)w_s}{N_0+1-n-m}\right] \end{aligned} \tag{7}$$

and the passband energy $E_p(h_0)$ with passband cut-off frequency w_p is

$$\begin{aligned} E_p(h_0) = \int_0^{w_p} (H_0(w) - 1)^2\,dw &= \left(h_0\left(\frac{N_0+1}{2}\right) - 1\right)^2 w_p \\ &+ 4h_0((N_0+1)/2 - 1)\sum_{n=1}^{(N_0-1)/2} h_0(n)\frac{\sin((N_0+1)/2 - n)w_p}{(N_0+1)/2 - n} \\ &+ 2\sum_{n=1}^{(N_0-1)/2} h_0(n)^2\left[w_p + \frac{\sin(N_0+1-2n)w_p}{N_0+1-2n}\right] \\ &+ 2\sum_{n=1}^{(N_0-1)/2}\sum_{m=1,m\neq n}^{(N_0-1)/2} h_0(n)h_0(m) \\ &\times \left[\frac{\sin(n-m)w_p}{n-m} + \frac{\sin(N_0+1-n-m)w_p}{N_0+1-n-m}\right] \end{aligned} \tag{8}$$

In the wavelet theory, wavelet properties such as regularity and smoothness of filters are used as design criteria [2,35,48]. Regularity refers to the convergence to a continuous function in iterating the wavelet filter. For a low-pass filter $H_0(z)$, the regularity of order m requires at least m zeros of its amplitude response $H_0(w)$ at $w = \pi$, i.e.

$$H_0(w)|_{w=\pi} = 0 \quad \text{and} \quad \frac{d^i H_0(w)}{dw^i}|_{w=\pi} = 0 \quad \text{for } i = 1, 2, \ldots, m-1. \tag{9}$$

Similarly, for a high-pass filter $H_1(z)$, the regularity of order m requires at least m zeros of its amplitude response $H_1(w)$ at $w = 0$.

It has been shown that regularity of order 2 is generally good enough for subband image coding [3,5]. Regularity of order 2 can be achieved by setting $H_0(w = \pi) = 0$ for the low-pass filter and $H_1(w = 0) = 0$ for the high-pass filter, which give us

$$\sum_{n=1}^{N_0} (-1)^{n+1} h_0(n) = 0 \quad \text{and} \quad \sum_{n=1}^{N_1} h_1(n) = 0. \tag{10}$$

Furthermore, the following condition is necessary in making infinite iteration of the low-pass and high-pass filter converges [2]

$$\sum_{n=1}^{N_0} h_0(n) = 1 \quad \text{and} \quad \sum_{n=1}^{N_1} (-1)^{n+1} h_1(n) = 1. \tag{11}$$

Although the concept of regularity is helpful, regularity by itself is not sufficient for choosing good filters for image compression. Therefore, other criteria such as the impulse response and step response of individual filters have been proposed [48]. So far, the sufficient condition of filter banks being optimal for image compression remains an open problem.

3. Optimization and generalization based design

In this section, we present a new method for designing filter banks for image compression based on sample images. The goal is to show that, with a good optimization and generalization procedure, it is possible to find coefficient sets that improve the compression quality of the best existing filter banks across multiple images, compression procedures, and compression ratios.

Since different types of images may have different characteristics, it may not be possible to group them together. Therefore, one approach is to treat individual sample images as separate problem domains and evaluate and optimize the performance of filter banks accordingly. In generalization, the performance of candidate filter banks can be estimated statistically based on the sample images.

Our new method has two major components: optimization and generalization. In the optimization phase, we formulate the design problem as a nonlinear optimization problem whose objective consists of the performance metrics of image compression and of the individual filters. Filter banks that optimize the design objectives are sought based on a set of training images. One distinctive feature of this approach is that several optimal or near-optimal coefficient sets are obtained based on individual training images. In the generalization phase, we select one generalizable coefficient set from the results obtained in the optimization phase. The final result performs well not only on the training examples used during the optimization process, but also on test cases not seen. This approach allows us to find the best filter coefficients in the context of a real image compression procedure for a class of images.

3.1. Optimization phase

Issues to be addressed in the optimization phase include problem formulation, performance evaluation, and selection of optimization methods. First, the design problem is formulated as an optimization problem to maximize image compression performance. Then, the performance of a filter bank coefficient set is evaluated by compressing training images using some image compression algorithm. This is the most time consuming part in the design method. Finally, nonlinear optimization methods are used to find optimal or near-optimal solutions of the optimization problem. In this subsection, we discuss the problem formulation issues in more detail.

The goal of the design method is to find filter coefficients in a biorthogonal filter bank that maximizes the image compression performance. Mathematically, the objective to be optimized is some combination of the performance metrics (shown in Table 1) of the overall image compression application, the overall filter bank, and the individual filters. In the design of biorthogonal filter banks, the lengths of filters are fixed according to the structure of the filter bank. Therefore, the filter coefficients are the only variables to be solved.

In the problem formulation, the following constraints are formulated based on the essential properties of biorthogonal filter banks to reduce the search space.

Constraint 3.1.1. The coefficients of the synthesis filters are chosen based on those of the analysis filters according to Eq. (2). By doing that, the design of filter banks becomes finding the coefficients of the analysis filters $H_0(z)$ and $H_1(z)$.

Constraint 3.1.2. The total length of the two filters is a multiple of 4.

Constraint 3.1.3. The two analysis filters are FIR filters of even length or odd length at the same time. When the two filters are of even length, filter $H_0(z)$ is symmetric and filter $H_1(z)$ is antisymmetric. When they are of odd length, both of them are symmetric.

Constraint 3.1.4. The two analysis filters satisfy the wavelet conditions stated in Eqs. (10) and (11).

The first three constraints are part of the requirements for the PR biorthogonal filter banks that we have discussed in Section 2.1. We relax the other requirements because it is unnecessary in the lossy compression situation to have the PR filter banks. The last constraint on individual analysis filters is a collection of linear equations. It can be used to remove variables through substitution to reduce the search space.

The PSNR and compression ratio are the most widely used performance metrics in image compression. The interaction between these two conflicting objectives gives rise to a set of compromised solutions, known as the Pareto-optimal solutions [8,42] that no one of them is better than any other with respect to both metrics. Previous approaches to deal with multi-objective problems can be divided into three classes, depending on the way that preferences are set: (a) *methods that set prior preferences*, which include forming a weighted sum of the objectives [26,33], constraining some objectives [19,49], and using reference points [43]; (b) *methods that change preferences interactively*, which include local gradient search [29,34] and iterative reduction of nondominated sets [47]; and (c) *methods that use posterior preferences*, which include evolutionary algorithms that may find Pareto-optimal solutions first before selecting preferred ones [13,17,41].

In our method, we apply the first approach by constraining the compression ratio and optimizing the PSNR. The goal is to find a filter bank that satisfies the constraint on the compression ratio and has the best performance in the unconstrained measure, PSNR. The optimization problem is therefore formulated as follows:

$$\max_{h_0,h_1} \ \mathrm{PSNR}(h_0,h_1,A,I,\mathrm{CR}) \tag{12}$$

$$\text{subject to } \sum_{n=1}^{N_0}(-1)^{n+1}h_0(n)=0, \quad \sum_{n=1}^{N_1}h_1(n)=0,$$

$$\sum_{n=1}^{N_0}h_0(n)=1, \quad \sum_{n=1}^{N_1}(-1)^{n+1}h_1(n)=1, \tag{13}$$

$$\text{Constraints 3.1.1, 3.1.2 and 3.1.3,}$$

where h_0 and h_1 are the coefficients of the analysis filters in the biorthogonal filter bank with length N_0 and N_1, respectively; A is the image compression

algorithm; I is the input image; CR is the compression ratio; and PSNR is a nonlinear function of h_0, h_1, A, I, and CR, and is measured based on the image compressing algorithm A, the input image I, and the degree of compression CR. The optimization problem is a nonlinear optimization problem with a highly nonlinear objective.

The optimization problem defined in (12) can be solved by various existing optimization methods, including multi-starts of local descents such as gradient-descent and quasi-Newton methods, simulated-annealing [15,27,40], and genetics-based methods [18,25]. Empirically, applying these methods from random starting points to solving the problem does not yield good results. The reason is that the majority portion of the search space corresponds to filter coefficients that lead to equally bad coding results and poor PSNR values. In another words, the search space is a big plateau dotted with small, steep hills. On the plateau, the objective function PSNR does not provide useful guidance for the search to progress towards the hills. To overcome this difficulty, we introduce an additional term to the objective function that is based on the stopband and passband energies of the individual filters. These energy functions provide the necessary guidance for the search to travel towards the hills on the plateau. The energies are calculated according to the formulas in Eqs. (7) and (8).

There are two ways to incorporate the additional objective term:

- The first approach is based on the subband coding theory and makes the analysis filters to approximate the ideal low-pass and high-pass filters, i.e., minimizing the total stopband and passband energies of the two analysis filters:

$$\min_{h_0,h_1} E(h_0,h_1) = E_s(h_0) + E_s(h_1) + E_p(h_0) + E_p(h_1). \tag{14}$$

  This objective implies that the smaller the energy the better. Unfortunately, this is not always the case in image compression. Although the filter banks that generate the best PSNR values tend to have small $E(h_0,h_1)$ values, they usually do not have the smallest.
- The second approach is to minimize the difference between the energy of the analysis filters (h_0,h_1) and that of the baseline solution (h_0^*,h_1^*), i.e.,

$$\min_{h_0,h_1}(E(h_0,h_1) - E(h_0^*,h_1^*))^2. \tag{15}$$

  This formulation uses the baseline solution as the reference point of the energy value. This objective implies that good solutions tend to have an energy level similar to that of the baseline solution. In our design method, the best existing filter banks are used as the baseline solutions. The formulation works well in our experiments.

In combining the second objective with the main objective in Eq. (12), we need to strike a right balance between them. A weight-sum formulation is used

to combine the maximization of PSNR and the minimization of filter energies as follows

$$\max_{h_0,h_1} w\mathrm{PSNR}(h_0,h_1,A,I,\mathrm{bpp}) - (1-w)(E(h_0,h_1) - E(h_0^*,h_1^*))^2, \tag{16}$$

where $0 \leqslant w \leqslant 1$ is a constant weight; h_0^* and h_1^* are the filters in the baseline filter bank. Note that when $w = 1$, the objective becomes maximizing PSNR.

3.2. Generalization phase

Generalization is an important step in the design process because only a small number of sample images are usually used during the optimization phase whereas the final filter banks designed are expected to work well on other images not seen in optimization. Generalization has commonly been used as a post-design verification phase that simply verifies the generalizability of the design results by evaluating them on a new set of sample images. This approach is suitable when sample images used in optimization are representative of all the images in the image domain. When sample images used in generalization have different characteristics, the design results may not be generalized.

To compare filter banks bearing different performance distribution across different images in an application domain, we need to have a performance metric that is independent of the actual distributions. In our approach, we use a new metric called *probability of win* to measure the probability that a particular filter bank is better than another filter bank [19,49,50]. Since the probabilities are between 0 and 1, we eliminate the dependence on actual performance distributions. Using this metric, we can verify whether a filter bank is generalizable across images of different performance distributions.

The statistical metric, *probability of win* (P_{win}), is used to compare the performance of filter banks in a range-independent and distribution-independent fashion. The probability of win measures statistically how much better (or worse) the sample mean of one hypothesis, μ_1, is as compared to that of another, μ_2. It resembles the significance level in general hypothesis testing, but there are two major differences. First, only one hypothesis $\{H : \mu_1 > \mu_2\}$ is specified, without the alternative hypothesis. Further, in contrast to hypothesis testing, acceptance confidence is not given in advance but is evaluated based on sample values. The advantage of P_{win} is that it considers both the mean and the variance of the performance data.

Under the assumption that the performance values of filter banks are independent and normally distributed, two filter banks, S_i and S_j, can be compared based on the probability of win. First, we calculate the difference of performance values between the two filter banks, $P_i - P_j$, on the N sample images. Assuming that the sample mean of $P_i - P_j$ is $\hat{\mu}$ and sample variance is $\hat{\sigma}^2$, then $P_{\mathrm{win}}(S_i, S_j)$ is defined as follows

$$P_{\text{win}}(S_i, S_j) = F_t\left(N-1, \frac{\hat{\mu}}{\sqrt{\hat{\sigma}^2/N}}\right), \tag{17}$$

where $F_t(v,x)$ is the cumulative distribution function of Student's t-distribution with v degrees of freedom, and P_{win} is the probability that the true performance (population mean) of S_1 is better than that of S_2. When $N \to \infty$, we have

$$P_{\text{win}}(S_i, S_j) \approx \Phi\left(\frac{\hat{\mu}}{\sqrt{\hat{\sigma}^2/N}}\right), \tag{18}$$

where Φ is the standard cumulative normal distribution function [14].

Based on P_{win}, the generalization is performed in our method as follows:

1. Assume the k filter banks found in the optimization phase are S_i $(i = 1, \ldots, k)$, and the baseline filter bank is S^*. Evaluate S_i $(i = 1, \ldots, k)$ and S^* on all m training images, I_j $(j = 1, \ldots, m)$ with the same compression ratio used in the optimization phase and obtain the corresponding PSNR values, $P_j(S_i)$ $(i = 1, \ldots, k)$ and $P_j(S^*)$.
2. Compute the improvement in the PSNR value for each filter bank over the baseline solution on each training image: $P_j(S_i) - P_j(S^*)$, $j = 1, \ldots, m$. Filter banks with an average PSNR improvement smaller than 0 over the baseline solution are discarded. This step improves the speed of the method by throwing away unpromising solutions.
3. Evaluate the remaining filter banks and the baseline solution on each training images at the other compression ratios under consideration, obtain the corresponding PSNR values, and compute the improvement of the filter banks in PSNR over the baseline solution.
4. For each filter bank S_i, calculate the P_{win} against the baseline solution, $P_{\text{win}}(S_i, S^*)$, The filter bank with the best P_{win} value is selected as the final result.

4. Agent-based system development

Agent-based approach is a new paradigm in developing open, distributed software systems and has been applied to a wide variety of applications ranging from small systems, such as information filtering, to large, open and complex systems, such as air traffic control [6,24,28,31]. Agents in multi-agent systems have different roles and work on different tasks [44,53,54]. Through communication, collaboration, and coordination, they jointly solve a complex problem. By using prevalent and standard technologies, such as

Java and COBRA, agent-based approach offer better portability and scalability. In view of the advantages, we adopt the agent-based approach in developing the prototype of the optimization and generalization based design method.

The design process consists of several inter-dependent components that are handled by specialized software agents. Specifically, the major components perform problem formulation, performance evaluation, optimization, and generalization. In the system, each component is handled by an software agent implemented using the Java Agent Template Lite (JATLite) libraries [23]. JATLite, written in Java, allows easy creation of software agents that communicate robustly over the Internet. An agent in the environment makes a single connection to the Agent Message Router (AMR) that forwards messages to other agents. JATLite provides a convenient template for building agents that utilize a common high-level language and protocol. The agent-based approach makes the software system modular, portable, and compatible with applications developed by multiple sources [37,38]. In addition, agents are naturally distributed, enabling the design to be done in parallel and distributed environments.

Fig. 2 shows a multi-agent framework for the optimization and generalization based filter bank design method. The system consists of four major software agents: formulation agent, image coding agent, optimization agent, and generalization agent. The software agents are specialized to performing their corresponding tasks. Jointly they search for filter banks that optimize the image compression quality based on training images. In this section we discuss the functionality of each agent in detail.

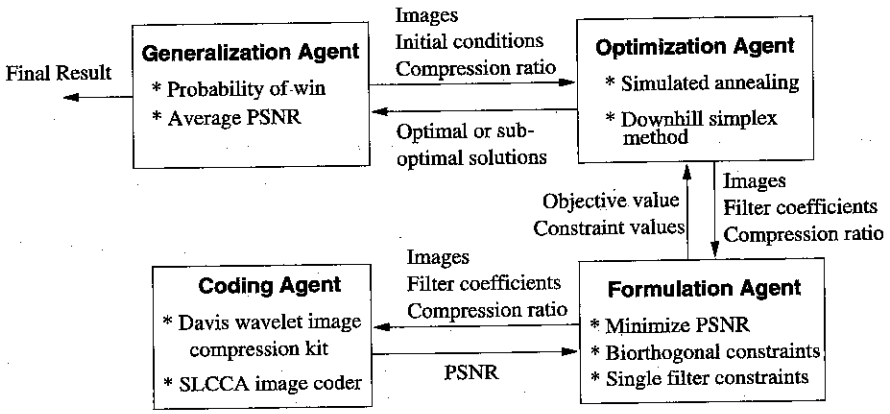

Fig. 2. A multi-agent framework for the optimization and generalization-based approach of filter bank design. Specialized agents in the system perform generalization, problem formulation, optimization, and image compression, concurrently.

4.1. Problem formulation agent

The formulation agent builds an optimization model for the filter bank design problem and uses it to evaluate the objective and constraint values for the optimization agent. In general, different optimization models can be created based on the properties of design variables (filter coefficients) and alternative design criteria and constraints. This enables the system to answer what-if questions and to explore alternatives. The construction of optimization models should also consider the optimization tools available in the system so that good solutions can be found efficiently in practical situations.

In the filter bank design application, the goal is to find filter banks that maximize the image compression quality. Two commonly used and inter-dependent performance metrics in image compression are the PSNR and the compression ratio. The PSNR in decibels (dB) is defined as

$$\mathrm{PSNR} = 20\log_{10}\frac{2^k - 1}{\mathrm{RMSE}}, \tag{19}$$

where k is the number of bits per pixel (bpp) in the image; and RMSE is the root mean-squared error (pixel-wise) between the original image and the reconstructed image. The compression ratio is defined as the number of bits of the original image divided by that of the compressed image.

In implementing the optimization phase of our method, we fix the compression ratio and search for filter banks that maximize the PSNR. The optimization formulation we have used consists of the objective function in Eq. (16) and the constraints in Eq. (13).

4.2. Image coding agent

The performance of a filter bank in image compression, such as the PSNR value at a certain compression ratio, is evaluated by the image coding agent based on an input image. We have used two image compression programs in our experiments: the SLCCA image coder [7] and the Davis wavelet image compression construction kit [12]. The Davis coder consists of invertible transform, quantization, and entropy coding. It applies simple strategies in each compression step and is easy to work with. The biorthogonal filter banks being designed are used in its transform step. The SLCCA is a high performance wavelet image coder that applies efficient representation of wavelet data and exploits both within-subband clustering of significant coefficients and cross-subband dependency in significant fields. It performs better than the Davis coder in compressing the sample images and is faster in each function evaluation. However, SLCCA requires manual tuning to reach a specified compression ratio, which is very time-consuming.

4.3. Optimization agent

The optimization agent applies various optimization methods to search for the best filter banks. It calls the formulation agent to get the necessary objective and constraint values. Multiple optimization agents may run simultaneously using different optimization models, different coding algorithms, and different sample images.

The optimization agent supports the integration of optimization models and optimizers. Based on the problem characteristics, such as the type of objective and constraint functions, the optimization agent applies the proper optimization method to solve it. The multi-agent system provides a flexible way to explore alternatives because both the optimization model as well as the number of variables, and the objective and constraint functions can be altered easily.

In the experiments of filter bank design, the optimization agent applies either the adaptive simulated annealing algorithm (ASA) [21,20] or multi-starts of the downhill simplex method (DSM). The ASA is an efficient implementation of the simulated annealing global optimization method. It has been applied to many applications and achieve good results. The DSM is a classical local search method that converges to a local optimum from an initial point. It has the benefit of fast convergence. However the result it converges to may be far away from the global optimal point.

4.4. Generalization agent

The generalization agent is responsible for decomposing tasks and distributing them to other agents in the beginning, and collecting and integrating the results in the end. For the filter bank design problem, multiple optimization tasks with different initial conditions such as the initial values of the filter coefficients, and different training images are distributed to multiple optimization agents. Each optimization agent finds an optimal filter bank based on the given image. After the optimization agents finish their jobs and obtain corresponding optimal filter banks, the generalization agent collects these results and selects one by applying the generalization procedure based on the probability of win metric. The final result should perform well not only on the training images, but also on other unseen images.

The run-time scenario of the multi-agent system is as follows. First, the generalization agent distributes multiple optimization tasks to different optimization agents. Then, each optimization agent works with its formulation agent and image coding agent, and applies the appropriate optimization method. In the end, the results are returned back to the generalization agent. Overall, the multi-agent system approach is modular, flexible, and suitable for parallel and distributed processing.

5. Experimental results

In our experiments, we have applied the optimization and generalization based method to design biorthogonal filter banks for image compression and obtained promising results. Our goal is to find generalizable filter banks that perform better than the best known solutions.

The benchmark images used in the experiments, shown in Fig. 3, consist of eight natural images and eight texture images. Three natural images ("Lena", "Barbara", "Goldhill") and three texture images ("Sweater", "Raffia", and "Water") were picked arbitrarily as the training images used in the optimization and generalization phases of our method. The other images were used to test the performance of the final result. In the experiments, the performance of filter banks was evaluated using the Davis wavelet image coder and the SLCCA image coder. In the design of 9/7 biorthogonal filter banks, the widely used Antonini's 9/7 filter bank [2] was used as the baseline solution.

In the optimization phase, either the multi-starts of DSM or the ASA was applied. The performance of filter banks was evaluated based on a fixed compression ratio, 32:1. In calculating the passband and stopband energies, the passband cut-off frequency w_p was set to 1 and the stopband cut-off frequency w_s was set to $\pi - 1$. In our experiments, the initial points of ASA were always the baseline solution. In the multi-starts of DSM, initial points were randomly generated inside a box region either around the baseline solution or around the origin 0. ASA was much slower than DSM, but usually found better solutions.

In the generalization phase, the objective is to choose the filter bank that not only performs well on the training images, but also generalizes well. Based on the filter banks obtained in the optimization phase, the generalization agent selects the best one through the following steps: (a) Evaluate the candidate filter banks and the baseline solution on all training images at compression ratios 16:1, 32:1, and 64:1, and obtain the corresponding PSNR values; (b) For each filter bank, compute the aggregated performance measure, P_{win}, against the baseline solution on all training images. The filter bank with the best P_{win} is selected as the final result.

In the first experiment of designing 9/7 biorthogonal filter banks, DSM was used as the optimization method and Eq. (16) with $w = 1$ was the objective. The compression procedure was the Davis coder. DSM was run 10 times from different random initial points around the origin 0 or the baseline solution, the Antonini's filters. Different ranges of the initial points (0.01, 0.05, 0.1, and 0.3) were tried in the experiments. The step size of DSM was set to 0.001. Each function evaluation using the Davis coder took 18 s for the natural images and 8 s for the texture images. It took more time in evaluating the natural images because they are four times as large as the texture images. All DSM runs converged within 500 function evaluations, which correspond to 2.5 h for a natural image and 1.1 h for a texture image.

Fig. 3. The 16 benchmark images used in our experiments. In the order from left to right and from top to bottom, they are "Lena", "Barbara", "Goldhill", "Baboon", "Man", "Couple", "Tank", "Boat", "Sweater", "Raffia", "Water", "Fingerprint", "Pig", "Wool", "Sand", and "Grass". The first eight are 512 × 512 pixel natural images. The last eight are 256 × 256 pixel texture images.

Table 2 shows the P_{win} of the best filter banks obtained using different training images in the optimization phase. The P_{win} is evaluated against the baseline solution based on the six training images or the total 16 images. The training images are used in two different ways in the optimization phase. One is to use individual training images in the performance evaluation during optimization and obtain an optimal filter bank for each training image. Another approach is to use a subset or the whole set of training images during optimization. The advantages of the first approach are fast performance evaluation because the time is proportional to the number of images used, and more diverse candidate filter banks to be used in the generalization phase. The results of the two approaches were compared in the experiments.

Table 2 shows that the solutions are generally not good when the initial points are around the origin 0. Occasionally, good solutions can still be found. The best result was obtained using the Sweater image and the initial range 0.1. Its P_{win} value is 0.880 on the six training images. When the initial points are around the baseline solution, the results generally get worse as the initial range is increased from 0.01 to 0.3. The results obtained using images Lena, Barbara, Goldhill, and Sweater are better than the baseline solution (P_{win} larger than 0.5), whereas those obtained using images Raffia and Water are worse. The best filter bank was found by using Goldhill with initial range 0.1. Its P_{win} values against the baseline solution are 0.931 and 0.975, on the six training images and on the total 16 images, respectively. In comparison, we also tried using the average performance on all six training images as the optimization objective. The best filter bank obtained is worse than the one obtained from Goldhill. Its P_{win} values are 0.901 and 0.772 based on the six training images and the total 16 images, respectively.

In the second experiment, ASA was used as the optimization method and Eq. (16) with $w = 1, 0.8, 0.6, 0.4$, and 0.2 were the objectives. The initial point was always the baseline solution. To speedup the convergence of ASA, we restricted the search ranges on the variables. In the design of 9/7 filter banks, the search ranges of the first three coefficients of $H_0(z)$ were limited to $[0.01, 0.1]$, $[-0.1, -0.01]$, and $[-0.12, -0.11]$, respectively. The search ranges of the first two coefficients of $H_1(z)$ were limited to $[0.01, 0.1]$ and $[-0.1, -0.01]$, respectively. The execution of ASA was limited to 10,000 function evaluations in each run, which corresponds to 50 h for a natural image and 22 h for a texture image.

Table 3 shows the P_{win} of the best filter banks obtained using different training images in the optimization phase. The P_{win} is evaluated against the baseline solution based on the six training images or the total 16 images. As the weight w is reduced from 1 to 0.2, corresponding to less weight on the PSNR and more weight on the energy term, the final results gradually got worse when individual training images are used in measuring the performance. This trend was not observed when the average performance on all training images is used

Table 2
The P_{win} of the best filter banks obtained using different training images in the optimization phase[a]

Image optimized	Range of the initial points: around the Antonini's filters							
	0.01		0.05		0.1		0.3	
	6 images	16 images	6 images	16 images	6 images	16 images	6 images	16 images
Lena	0.901	0.974	0.085	0.013	0.104	0.090	0.059	0.101
Barbara	0.885	0.899	0.527	0.647	0.097	0.093	0.002	0.000
Goldhill	0.877	0.939	0.574	0.822	0.931	0.975	0.381	0.523
Sweater	0.809	0.849	0.402	0.408	0.076	0.025	0.001	0.001
Raffia	0.089	0.089	0.073	0.022	0.022	0.006	0.004	0.014
Water	0.392	0.476	0.399	0.346	0.010	0.0018	0.013	0.002
6 images	0.901	0.772	0.736	0.734	0.093	0.009	0.002	0.001
	Range of the initial points: around the origin 0							
Lena	0.000	0.000	0.000	0.000	0.007	0.010	0.001	0.000
Barbara	0.000	0.000	0.001	0.001	0.013	0.004	0.556	0.571
Goldhill	0.079	0.073	0.550	0.754	0.001	0.000	0.343	0.536
Sweater	0.000	0.000	0.003	0.000	0.880	0.972	0.015	0.015
Raffia	0.001	0.000	0.002	0.000	0.031	0.007	0.006	0.002
Water	0.009	0.000	0.004	0.002	0.191	0.042	0.077	0.009
6 images	0.000	0.000	0.005	0.007	0.016	0.003	0.330	0.198

[a] The optimization method is DSM and the compression procedure is the Davis coder. The initial points are randomly generated around the baseline solution, the Antonini's filters, or around the origin 0. The ranges for generating the random initial points are 0.01, 0.05, 0.1 and 0.3. The P_{win} are evaluated based on the six training images and the total 16 images.

Table 3
The P_{win} of the best filter banks obtained using different training images in the optimization phase[a]

Image optimized	Weight w in the objective function, Eq. (16)									
	1		0.8		0.6		0.4		0.2	
	6 images	16 images	6 images	16 images	6 images	16 images	6 images	16 images	6 images	16 images
Lena	0.830	0.724	0.570	0.285	0.564	0.423	0.652	0.324	0.134	0.029
Barbara	0.831	0.959	0.709	0.900	0.422	0.784	0.835	0.954	0.709	0.902
Goldhill	0.873	0.955	0.889	0.940	0.702	0.942	0.621	0.914	0.500	0.140
Sweater	0.368	0.187	0.152	0.100	0.339	0.579	0.233	0.169	0.208	0.094
Raffia	0.120	0.025	0.112	0.022	0.130	0.027	0.102	0.020	0.063	0.018
Water	0.303	0.218	0.303	0.218	0.296	0.195	0.415	0.644	0.065	0.090
6 images	0.893	0.959	0.893	0.960	0.879	0.966	0.902	0.963	0.919	0.954

[a] The optimization method is ASA and the compression procedure is the Davis coder. The initial point is the baseline solution, the Antonini's filters. The P_{win} are evaluated based on the six training images and the total 16 images.

to compute the objective values. The best filter bank was obtained with $w = 0.2$ by using the average performance on all six training images in calculating the optimization performance. Its P_{win} values are 0.919 and 0.954, based on the six training images and the total 16 images, respectively. The best filter bank obtained using individual training images in optimization is from Goldhill with $w = 0.8$. Its P_{win} values are 0.889 and 0.940, based on the six training images and the total 16 images, respectively.

Finally, in the last experiment, the SLCCA was used as the compression procedure and the optimization method was ASA. Although SLCCA is faster than the Davis coder in each function evaluation (5 s by SLCCA vs. 18 s by Davis for a natural image), it needs time-consuming manual tuning to get to the right compression ratio. Thus a limited runs were tried in the experiment. The objective was Eq. (16) with $w = 0.2$. Three images, Lena, Barbara, and Goldhill, were used as the individual images to be optimized. The initial point of ASA was always the baseline solution. The final results are shown in Table 4. The best filter bank was obtained using Goldhill. Its P_{win} values are 0.911 and 0.977 based on the six training images and the total 16 images, respectively.

Table 5 summarizes the performance of the best results obtained in the experiments, DSM + Davis, ASA + Davis, and ASA + SLCCA. Both the P_{win} values and the average improvement of *PSNR* values on the six training images and the total 16 images are shown in the table. The same filter banks are tested in the two image coders. They all improve the baseline solution (the Antonini's filter bank) on both performance measures, P_{win} and PSNR. Although the designs were based on the six training images and the Davis coder, the results work well on the total 16 images. Using the Davis image coder, they have P_{win} values at least 0.954 and the average improvement in PSNR from 0.047 to 0.065 on the total 16 images. Using the SLCCA coder, they have P_{win} values at least 0.964 and the average improvement in PSNR from 0.027 to 0.051.

Table 6 shows the performance comparison of the new 9/7 filter bank designed using ASA + SLCCA and the Antonini's filter bank. The performance (PSNR) on the 16 individual images is compared on three compression ratios, 16:1, 32:1, and 64:1, and using two image coders, Davis and SLCCA. The data

Table 4
The P_{win} of the best filter banks obtained using different training images in the optimization phase[a]

Image optimized	Image tested	
	6 images	16 images
Lena	0.874	0.964
Barbara	0.501	0.410
Goldhill	0.911	0.977

[a] The optimization method is ASA and the compression procedure is SLCCA. The P_{win} are evaluated based on the six training images and the total 16 images.

Table 5
Performance summary of the best results obtained in the experiments, DSM + Davis, ASA + Davis, and ASA + SLCCA[a]

Design method	Evaluation method							
	Davis image coder				SLCCA image coder			
	6 images		16 images		6 images		16 images	
	P_{win}	$PSNR_{impr}$	P_{win}	$PSNR_{impr}$	P_{win}	$PSNR_{impr}$	P_{win}	$PSNR_{impr}$
DSM + Davis	0.931	0.070	0.975	0.047	0.833	0.033	0.968	0.027
ASA + Davis	0.919	0.082	0.954	0.054	0.727	0.035	0.964	0.032
ASA + SLCCA	0.911	0.090	0.977	0.065	0.790	0.059	0.967	0.051

[a] Both the P_{win} values and the average improvement of PSNR values on the six training images and the total 16 images are shown in the table. Both the Davis and SLCCA image coders are used in the evaluation.

Table 6
Performance comparison (PSNR [dB]) of the Antonini's 9/7 filter bank (A) and the new 9/7 filter bank (B) designed using ASA + SLCCA by the proposed method on the 16 benchmark image based on the SLCCA image coder [7] and the Davis image coder [12], respectively

Image	Compression ratio					
	16:1		32:1		64:1	
	A	Impr. of B	A	Impr. of B	A	Impr. of B
SLCCA image coder						
Lena	37.35	−0.01	34.28	+0.05	31.25	+0.07
Barbara	31.89	+0.22	28.18	+0.16	25.36	+0.06
Goldhill	33.26	+0.01	30.65	+0.02	28.61	0.00
Baboon	25.86	+0.02	23.45	+0.03	21.88	+0.01
Man	33.19	−0.01	30.10	0.00	27.73	−0.01
Couple	32.59	−0.01	29.35	−0.02	26.96	0.00
Boat	34.61	+0.02	31.11	0.00	28.25	+0.01
Tank	31.27	+0.04	29.44	+0.04	27.95	+0.02
Sweater	43.67	−0.25	38.70	+0.02	35.31	+0.05
Raffia	21.97	+0.23	18.49	+0.09	16.02	+0.03
Water	30.49	+0.16	27.39	+0.11	24.55	+0.04
Fingerprt	28.57	+0.22	25.04	+0.16	22.26	+0.18
Pig	27.67	+0.08	25.01	+0.04	23.19	+0.01
Wool	27.11	+0.12	24.33	+0.08	22.37	+0.02
Sand	25.03	+0.14	22.32	+0.05	20.34	+0.01
Grass	26.38	+0.12	23.29	+0.04	21.19	−0.03
Davis image coder						
Lena	36.18	+0.05	33.17	−0.07	30.23	+0.04
Barbara	29.54	+0.67	26.65	+0.18	24.31	+0.07
Goldhill	32.62	0.00	30.08	+0.04	28.22	+0.01
Baboon	25.05	+0.08	22.87	+0.01	21.62	+0.01
Man	32.29	+0.01	29.45	−0.05	27.36	+0.03
Couple	31.47	+0.07	28.52	+0.12	26.55	+0.04
Tank	31.09	+0.05	29.21	+0.06	27.93	−0.08
Boat	33.41	+0.06	30.14	+0.02	27.56	+0.04
Sweater	43.19	−0.12	38.39	+0.03	35.26	+0.07
Raffia	21.69	+0.28	17.99	+0.06	15.91	+0.02
Water	29.93	+0.11	27.06	+0.11	24.02	+0.11
Fingerprt	28.27	+0.15	24.66	+0.11	21.83	+0.15
Pig	27.33	+0.10	24.78	+0.02	22.92	−0.02
Wool	26.69	+0.13	24.04	+0.05	22.34	+0.07
Sand	24.75	+0.16	21.97	+0.01	20.14	+0.03
Grass	25.87	+0.08	23.01	+0.04	21.04	+0.01

show that the new filter bank improves the Antonini's filter bank for the majority of the images using both image coders and at the three compression ratios. For some images such as barbara and "Fingerprint", the improvement is significant. For instance, the improvement in PSNR is 0.67 dB for compressing ratio 16:1 using the Davis image coder on Barbara. Although the new

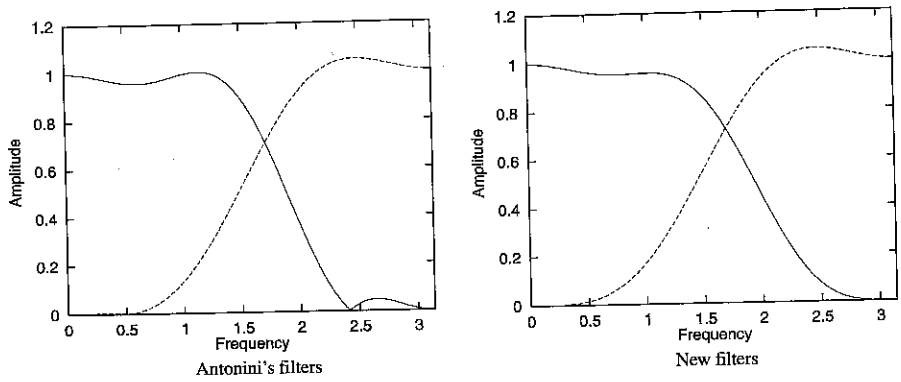

Fig. 4. Comparison of amplitude responses of the pair of analysis filters in the Antonini's (9,7) filter bank and those in the new (9,7) filter bank found by our new method.

filter bank was designed based on a small number of training images, one image coder, and a fixed compression ratio, it performs well on other images, using other image coder, and at different compression ratios.

Fig. 4 compares the amplitude responses of the two analysis filters in the new 9/7 filter bank designed using ASA + SLCCA and those of the Antonini's 9/7 filter bank. The Antonini's filters have high-order regularity and their amplitude responses are smoother. In contrast, the new filters have sharper transitions.

In addition to the design of 9/7 biorthogonal filter banks, we have applied our method to design 10/18 biorthogonal filter banks, trying to improve the widely used 10/18 filter bank published in [48]. We have successfully found filter banks better than the existing solution. However, the improvement is not as much as the results for the 9/7 filter banks.

6. Summary

In this paper, we present a new optimization and generalization-based method for designing filter banks in subband image coding. We formulate the design problem as a nonlinear optimization problem and focus on improving the performance of the best existing solutions. Our approach allows us to find filter banks in the context of image coding algorithms to maximize coding quality. Generalizable solutions that perform well not only on the training image used in optimization, but also on other images are chosen as the final result. In our multi-agent implementation, agents collaborate with each other to find better filter banks that improve image compress quality. In our experiments, we have found new 9/7 filter banks that improve the coding quality

of the Antonini's 9/7 filter bank, a widely used wavelet filter bank, in terms of PSNR on a set of benchmark images. Significant improvement has been obtained on some images using the two image coders. Experiment results show that optimization based on a single image followed by generalization is efficient and can achieve good results. Even though the new filter banks are designed based on a small number of training images, a particular image coder, and a fixed compression ratio, they perform well on other test images, using other image coder, and at different compression ratios. Our method is general and modular, and can be applied to the design of other types of filter banks such as multi-rate and multi-band filter banks.

Acknowledgements

Research supported in part by the National Science Foundation under grants MIP-9632316 and DUE-9980375 and Research Board Grant of University of Missouri System.

References

[1] A.N. Akansu, R.A. Haddad, Multiresolution Signal Decomposition, Academic Press, San Diego, CA, 1992.
[2] M. Antonini, M. Barlaud, P. Mathieu, I. Daubechies, Image coding using wavelet transform, IEEE Trans. Image Process. 1 (2) (1992) 205–220.
[3] I. Balasingham, T.A. Ramstad, On the relevance of the regularity constraint in subband image coding, in: Proc. 31st Asilomar Conference on Signals, Systems, and Computers, 1997.
[4] I. Balasingham, T.A. Ramstad, J.M. Lervik, Survey of odd and even length filters in tree-structured filter banks for subband image compression, in: Proc. Int. Conf. on Acoustics, Speech, and Signal Process., 1997, pp. 3073–3076.
[5] Ilangko Balasingham, On optimal perfect reconstruction filter banks for image compression, Ph.D. Thesis, Norwegian University of Science and Technology, 1998.
[6] J.M. Bradshaw, Software Agents, AAAI Press, The MIT Press, Cambridge, MA, 1997.
[7] B.B. Chai, J. Vass, X. Zhuang, Significance-linked connected component analysis for wavelet image coding, IEEE Trans. Image Process. 8 (6) (1999) 774–784.
[8] V. Chankong, Y.Y. Haimes, Multiobjective Decision Making Theory and Methodology, North-Holland, New York, 1983.
[9] C.-K. Chen, J.-H. Lee, Design of quadrature mirror filters with linear phase in the frequency domain, IEEE Trans. Circuits Syst. – II 39 (9) (1992) 593–605.
[10] C.D. Creusere, S.K. Mitra, A simple method for designing high-quality prototype filters for m-band pseudo QMF banks, IEEE Trans. Signal Process. 43 (4) (1995) 1005–1007.
[11] G. Davis, S. Chawla, Image coding using optimized significance tree quantization, in: Proc. Data Compression Conf., 1997, pp. 387–396.
[12] G.M. Davis. Wavelet image compression construction kit, http://www.cs.dartmouth.edu/gdavis/wavelet/wavelet.html, 1997.
[13] K. Deb. Evolutionary algorithms for multi-criterion optimization in engineering design, in: Proc. Evolutionary Algorithms in Engineering and Computer Science, 1999.

[14] J.L. Devore, Probability and Statistics for Engineering and the Sciences, Brooks/Cole, Monterey, CA, 1982.
[15] J. Engel, Teaching feedforward neural networks by simulating annealing, Complex Syst. 2 (1988) 641–648.
[16] N.J. Fliege, Multirate Digital Signal Processing, Wiey, New York, 1994.
[17] J. Horn, N. Nafpliotis, D.E. Goldberg, A niched Pareto genetic algorithm for multiobjective optimization, in: IEEE Symp. Circuits and Systems, 1991, pp. 2264–2267.
[18] B.R. Horng, A.N. Willon Jr., Lagrange multiplier approaches to the design of two-channel perfect-reconstruction linear-phase FIR filter banks, IEEE Trans. Signal Process. 40 (2) (1992) 364–374.
[19] A. Ieumwananonthachai, B.W. Wah, Statistical generalization of performance-related heuristics for knowledge-lean applications, Int. J. Artificial Intelligence Tools 5 (1-2) (1996) 61–79.
[20] A.L. Ingber, Adaptive simulated annealing (ASA): lessons learned, J. Control Cybern. 25 (1) (1996) 33–54.
[21] L. Ingber, Adaptive simulated annealing (ASA), http://www.ingber.com/.
[22] V.K. Jain, R.E. Crochiere, Quadrature mirror filter design in the time domain, IEEE Trans. Acoust. Speech Signal Process. 32 (2) (1984) 353–361.
[23] JATLite, http://java.stanford.edu/java_agent/html.
[24] N.R. Jennings, M.J. Wooldridge, Agent Technology: Foundations, Applications, and Markets, Springer, Berlin, 1998.
[25] J.D. Johnston, A filter family designed for use in quadrature mirror filter banks, in: Proc. Int. Conf. ASSP, 1980, pp. 291–294.
[26] P.P. Khargonekar, M.A. Rotea, Multiple objective optimal control of linear systems: the quadratic norm case, IEEE Trans. Autom. Control 36 (1) (1991) 14–24.
[27] S. Kirkpatrick Jr., C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, Science 220 (4598) (1983) 671–680.
[28] P. Maes, Modeling adaptive automonous agents, Artificial Life J. 1 (1994).
[29] M. Mollaghasemi, G.W. Evans, Multicriteria design of manufacturing systems through simulation optimization, IEEE Trans. Systems, Man, Cybern. 24 (9) (1994) 1407–1411.
[30] T.Q. Nguyen, P.P. Vaidyanathan, Two-channel perfect-reconstruction FIR QMF structures which yield linear-phase analysis and synthesis filters, IEEE Trans. Acoust. Speech Signal Process. 37 (5) (1989) 676–690.
[31] H.S. Nwana, N. Azarmi (Eds.), Software Agents and Soft Computing: Towards Enhancing Machine Intelligence, Springer, Berlin, 1997.
[32] J.E. Odegard, C.S. Burrus, Smooth biorthogonal wavelets for applications in image compression, in: Proc. DSP Workshop, Loen, Norway, 1996.
[33] J. Radecki, J. Konrad, E. Dubois, Design of multidimensional finite-wordlength FIR and IIR filters by simulated annealing, IEEE Trans. Circuits Syst. – II 42 (6) (1995) 424–431.
[34] J.L. Ringuest, T.R. Gulledge, An interactive multi-objective gradient search, Oper. Res. Lett. 12 (1992) 53–58.
[35] O. Rioul, Simple regularity criteria for subdivision schemes, SIAM J. Math Anal. 23 (1992) 1544–1576.
[36] A. Said, W.A. Pearlman, A new, fast, and efficient image codec based on set partitioning in hierarchical tress, IEEE Trans. Circuits Syst. Video Technol. 6 (1996) 243–250.
[37] Y. Shang, L. Li, Intelligent agents for designing filter banks in image compression. in: Proc. Parallel and Distributed Methods for Image Processing, SPIE Int. Symp. Optical Science, Engineering and Instrumentation, July 1999, pp. 98–107.
[38] Y. Shang, H. Shi, A web-based multi-agent system for interpreting medical images, World Wide Web 2 (4) (1999) 209–218.
[39] J.M. Shapiro, Embedded image coding using zerotrees of wavelet coefficients, IEEE Trans. Signal Process. 41 (1993).

[40] S. Shekhar, M.B. Amin, Generalization by neural networks, IEEE Trans. Knowledge Data Eng. 4 (2) (1992) 177–185.
[41] A.A. Song, A. Mathur, K.R. Pattipati, Design of process parameters using robust design techniques and multiple criteria optimization, IEEE Trans. Systems, Man, Cybern. 25 (11) (1995) 1437–1446.
[42] R.E. Steuer, Multiple Criteria Optimization: Theory, Computation, and Application, Wiley, New York, 1986.
[43] R.E. Steuer, Multiple Criteria Optimization: Theory, Computation and Application, Krieger Publishing Company, 1989.
[44] K. Sycara, K. Decker, A. Pannu, M. Williamson, D. Zeng, Distributed intelligent agents, IEEE Expert, 1996.
[45] T.E. Tuncer, T.Q. Nguyen, General analysis of two-band QMF banks, IEEE Trans. Signal Process. 43 (2) (1995) 544–548.
[46] P.P. Vaidyanathan, Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial, Proc. IEEE 78 (1) (1990) 56–93.
[47] D. Vanderpooten, P. Vincke, Description and analysis of some representative interactive multicriteria procedures, in: C. de Bruyn, G. Colson, E.Y. Rodin (Eds.), Models and Methods in Multiple Criteria Decision Making, Pergamon Press, Oxford, 1989, pp. 1221–1238.
[48] J.D. Villasenor, B. Belzer, J. Liao, Wavelet filter evaluation for image compression, IEEE Trans. Image Process. 2 (1995) 1053–1060.
[49] B.W. Wah, A. Ieumwananonthachai, L.C. Chu, A. Aizawa, Genetics-based learning of new heuristics: rational scheduling of experiments and generalization, IEEE Trans. Knowledge Data Eng. 7 (5) (1995) 763–785.
[50] B.W. Wah, A. Ieumwananonthachai, Shu Yao, Ting Yu, Statistical generalization: theory and applications (plenary address), in: Proc. Int. Conf. Computer Design, IEEE, Austin, TX, October 1995, pp. 4–10.
[51] B.W. Wah, Y. Shang, T. Wang, T. Yu, Global optimization of QMF filter-bank design using NOVEL, in: Proc. Int. Conf. Acoustics, Speech Signal Process., vol. 3, IEEE, April 1997, pp. 2081–2084.
[52] B.W. Wah, Y. Shang, Z. Wu, Discrete Lagrangian method for optimizing the design of multiplierless QMF filter banks, in: Proc. Int. Conf. Application Specific Array Processors, IEEE, July 1997, pp. 529–538.
[53] G. Weib, S. Sandip, Adaption and learning in multi-agent systems, in: Lecture Notes in Artificial Intelligence, vol. 1042, Springer, Berlin, 1996.
[54] G. Weiss (Ed.), Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence, The MIT Press, Cambridge, MA, 1999.
[55] J.W. Woods (Ed.), Subband Image Coding, Kluwer Academic Publishers, Dordrecht, 1991.