Int. J. Appl. Math. Comput. Sci., 2004, Vol. 14, No. 1, 105–111

GRADIENT FLOW OPTIMIZATION FOR REDUCING BLOCKING EFFECTS OF TRANSFORM CODING

FENG GAO*, XIAOKUN LI*, XUN WANG*, WILLIAM G. WEE*

* Department of Electrical and Computer Engineering and Computer Science
University of Cincinnati
e-mail: fgao2002@yahoo.com

This paper addresses the problem of reducing blocking effects in transform coding. A novel optimization approach using the gradient flow is proposed. Using some properties of the gradient flow on a manifold, an optimized filter design method for reducing the blocking effects is presented. Based on this method, an image reconstruction algorithm is derived. The algorithm maintains the fidelity of images while reducing the blocking effects. Experimental tests demonstrate that the presented algorithm is effective.

Keywords: gradient flow, blocking effects, optimization, transform coding

1. Introduction

The discrete cosine transform (DCT) plays an important role in image and video compression techniques. With the advancement of video communication, DCT has attracted even further attention. The International Standard Organization (ISO) uses it as a standard component for image and video compression in JPEG and MPEG (ISO, 1991; 1993). It is well known that the DCT has two main advantages. The first advantage is the feature of its excellent energy compaction for highly correlated data. It has been shown that the DCT is very close to the Karhunan-Loeve transform for first-order statistic Markov processes which can be used to model most digital images in communication (Jain, 1989). The second advantage is the fact that the computation of the transform is efficient. A fast DCT is available as in the fast discrete Fourier transform computation.

According to transform coding theory and some standards, a given image is divided into small $p \times q$ rectangular blocks. Generally, the blocks are chosen to be square, that is, $p = q$, and we denote the size of each block as $B \times B$ in this paper. The processing of the DCT on a block is known as the block discrete cosine transform (BDCT). The process of partitioning an entire image into blocks provides efficient hardware design and reduced computation time. However, since the BDCT is used block by block without considering the correlations between any two neighboring blocks, it results in block artifacts which appear on many edges between two neighboring blocks. This phenomenon is known as blocking effects. It deteriorates the quality of the decoded image. The blocking effects are encountered when the bit rate is further reduced, as in the case of a higher compression.

The research of methodologies for reduction of blocking effects has attracted much attention since the 1980s. In (Reeve and Lim, 1984), two methods, the overlap method and the filtering method, were proposed. These methods share the same advantage of simplicity in computation, but some disadvantages exist in both the methods. Since then many papers have been published on this research. Yang *et al.* (1993) presented two other methods, one using projection onto constrained convex sets to reconstruct decoded images, and the other using a constrained least-squares method with a high frequency filter to recover images. This seminal paper introduced the optimization idea into the problem of blocking effects. In turn, the paper (Yang *et al.*, 1995) offered an adaptive method for this problem. Local statistical properties and human perception were first introduced in this research. In the paper (Minami and Zakhor, 1995) the use of correlations between the intensity values of boundary pixels of neighboring blocks was presented to reduce the blocking effects. More recently, the paper (Kim *et al.*, 2000) introduced a recognition method used in (Won and Derin, 1992) to reduce blocking effects. The paper (Kim *et al.*, 2000) set forth a restoration filter design method using edge direction information, a constrained least-squares filter and classification with a model fitting criterion.

In this paper, we propose a novel approach to the problem of reducing blocking effects. Based on the fact that the location of all block boundaries is known, and only the pixel values on the block boundaries need to be smoothed, an optimal filter design method is presented on a constraint manifold. The constraint manifold can be regarded as a lower dimensional manifold imbedded into an N^2-dimensional linear space $\mathbb{R}^{N^2}$. So our research can be converted into an optimization problem on the constraint manifold. This problem can be solved by using the gradient flow method on the manifold. Based on this idea, two algebraic differential equations for optimal filter design and optimal reconstruction are proposed for the reduction of blocking effects. An algorithm based on algebraic differential equations is derived. Two experiments are given as a test of the proposed algorithm. These experiments demonstrate that the algorithm is effective.

The organization of this paper is as follows: Following Introduction, a mathematical description of blocking effects is presented in Section 2. Section 3 discusses design methods of optimal filtering and reconstruction of decoded images using gradient flow optimization on a manifold. Section 4 describes the proposed algorithm and two experiments as tests. The conclusion is included in Section 5.

2. Mathematical Description of Blocking Effects

After the BDCT transform, a decoded $N \times N$ image X with blocking effects can be expressed in a submatrix form as

$$X = \begin{pmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,n} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,n} \\ \cdots & & & \\ X_{n,1} & X_{n,2} & \cdots & X_{n,n} \end{pmatrix}, \tag{1}$$

where $X_{i,j}$ is a $B \times B$ submatrix, $i, j = 1, 2, \ldots, n$, and $n = N/B$ is an integer. Every $X_{i,j}$ is called a block. There exist blocking artifacts between every adjacent block boundaries. Such artifacts are called blocking effects.

Let $f^r_{i,j}$ and $f^l_{i,j}$ be respectively the last and first columns of the submatrix $X_{i,j}$ for every i, j and write

$$f^r_j = ((f^r_{1,j})^T, (f^r_{2,j})^T, \ldots, (f^r_{n,j})^T)^T,$$

$$f^l_j = ((f^l_{1,j})^T, (f^l_{2,j})^T, \ldots, (f^l_{n,j})^T)^T,$$

where T denotes the transpose. Then the difference vector $f^r_j - f^l_{j+1}$ is a measure of the blocking effects in the column direction of X. Define the column edge difference vector f_{ced} as

$$f_{ced} = ((f^r_1 - f^l_2)^T, (f^r_2 - f^l_3)^T, \cdots, (f^r_{n-1} - f^l_n)^T)^T,$$

whose norm $\|f_{ced}\|$ can be used to measure all blocking effects in the column direction.

In the same manner, let $g^t_{i,j}$ and $g^b_{i,j}$ be the first and last rows of the submatrix $X_{i,j}$, respectively. Write

$$g^t_j = (g^t_{j,1}, g^t_{j,2}, \ldots, g^t_{j,n}),$$

$$g^b_j = (g^b_{j,1}, g^b_{j,2}, \ldots, g^b_{j,n}).$$

Then the difference $g^b_j - g^t_{j+1}$ is a measure of the blocking effects in the row direction of X. Define the row edge difference vector f_{red} as

$$f_{red} = (g^b_1 - g^t_2, g^b_2 - g^t_3, \ldots, g^b_{n-1} - g^t_n)^T.$$

The norm of f_{red} can be used to measure all blocking effects in the row direction.

Let f be the N^2-dimensional vector composed of all columns of the decoded image matrix X. The elements of f are arranged such that the first N elements form the first column of X, the next N elements form the second column of X, and so on. Here f is called the image vector of X. It is easy to design two matrices R_c and R_r such that

$$R_c f = f_{ced}, \quad R_r f = f_{red}. \tag{2}$$

For the image matrix X, we can also define the corresponding block edge vector f_e as

$$f_e = \big((f^r_1)^T, (f^l_2)^T, (f^r_2)^T, \ldots, (f^r_{n-1})^T, (f^l_n)^T, g^b_1, g^t_2, \ldots, g^b_{n-1}, g^t_n\big)^T.$$

It is also easy to design two matrices Q_c and Q_r such that

$$Q_c f_e = f_{ced}, \quad Q_r f_e = f_{red}. \tag{3}$$

The image vector f and the matrices R_c and R_r will be used for optimal filter design for the decoded image X. The block edge vector f_e and the matrices Q_c and Q_r will be used for the optimal reconstruction of the decoded image X.

3. Optimal Reconstruction Design Method

Generally, from Section 2, we can see that f_{ced} and f_{red} provide all the information of edge differences between any two neighboring blocks of the decoded image matrix X. Therefore, $\|f_{ced}\|$ and $\|f_{red}\|$ can be used to measure the blocking effects. The larger $\|f_{red}\|$ and $\|f_{ced}\|$, the greater the blocking effects. Here, we propose a method to design an optimal filter H such that, when the image vector f of X passes through the filter H, the corresponding edge differences $\|\hat{f}_{ced}\| \stackrel{\text{def}}{=} \|R_c H f\|$ and $\|\hat{f}_{red}\| \stackrel{\text{def}}{=} \|R_r H f\|$ of the new image vector $\hat{f} \stackrel{\text{def}}{=} Hf$ can be kept to two given real parameters ϵ_1 and ϵ_2, respectively, that is

$$\|\hat{f}_{ced}\| = \epsilon_1, \quad \|\hat{f}_{red}\| = \epsilon_2. \tag{4}$$

Note that (4) constitutes a lower manifold in the linear space $\mathbb{R}^{N^2}$, which we call the constraint manifold. Therefore, the problem of reducing blocking effects is converted into the optimization problem: Design an optimal filter H on the constraint manifold. We then solve this problem using the gradient flow optimization method on the manifold hereafter.

Given a decoded image X with blocking effects, let f be its image vector. We design an optimal matrix filter H such that the new image vector $\hat{f} = Hf$ is close to the old image vector f with the property of making the block boundaries smooth and improving the quality of the decoded image X. It is expected that once H is designed, the new image vector $\hat{f}$ is obtained, and the new reconstructed image $\hat{X}$ is close to the old decoded image X with an improved peak signal to noise ratio (PSNR).

We can formulate the above idea as a typical optimization problem: Given a decoded image X and f as its corresponding image vector, find a matrix filter H such that

$$\min_{H} \|Hf - f\|^2 \tag{5}$$

subject to

$$\|R_c H f\|^2 = \epsilon_1^2, \tag{6}$$

$$\|R_r H f\|^2 = \epsilon_2^2, \tag{7}$$

where R_c and R_r are the same matrices as in (2), and ϵ_1 and ϵ_2 should be chosen properly so that the original image information is retained and the blocking effects are reduced.

Note that the constraints in the above formulation mean that for the new image vector $\hat{f} = Hf$, its column edge difference vector $\hat{f}_{ced}$ ($\hat{f}_{ced} = R_c H f$) and row edge difference vector $\hat{f}_{red}$ ($\hat{f}_{red} = R_r H f$) must satisfy some smoothness conditions.

Let us start solving the problem (5)–(7). Define

$$\begin{aligned}\psi_1(H) &= (Hf - f)^T (Hf - f)\\ &= \operatorname{tr}[(Hf - f)(Hf - f)^T],\\ \psi_2(H) &= (R_c H f)^T (R_c H f)\\ &= \operatorname{tr}[(R_c H f)(R_c H f)^T],\\ \psi_3(H) &= (R_r H f)^T (R_r H f)\\ &= \operatorname{tr}[(R_r H f)(R_r H f)^T],\end{aligned}$$

where $\operatorname{tr}(X)$ denotes the trace of X.

We have

$$\frac{\partial \psi_1(H)}{\partial H} = 2(Hf - f)f^T = 2Hff^T - 2ff^T, \tag{8}$$

$$\begin{aligned}\frac{\partial \psi_2(H)}{\partial H} &= R_c^T R_c H f f^T + R_c^T R_c H f f^T\\ &= 2R_c^T R_c H f f^T,\end{aligned} \tag{9}$$

$$\frac{\partial \psi_3(H)}{\partial H} = 2R_r^T R_r H f f^T. \tag{10}$$

Let $\psi(H) = \psi_1(H) + \lambda_1 \psi_2(H) + \lambda_2 \psi_3(H)$, where λ_1 and λ_2 are indeterminates. We have

$$\begin{aligned}\frac{\partial \psi(H)}{\partial H} &= \frac{\partial \psi_1(H)}{\partial H} + \lambda_1 \frac{\partial \psi_2(H)}{\partial H} + \lambda_2 \frac{\partial \psi_3(H)}{\partial H}\\ &= 2(H - I + \lambda_1 R_c^T R_c H + \lambda_2 R_r^T R_r H) f f^T,\end{aligned}$$

where I is the identity matrix of the appropriate dimensions.

So we can take the gradient flow of H as

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -2(H - I + \lambda_1 R_c^T R_c H + \lambda_2 R_r^T R_r H) f f^T. \tag{11}$$

Next, let us determine what conditions λ_1 and λ_2 satisfy.

Since H satisfies $\psi_2(H) = \epsilon_1^2$ and $\psi_3(H) = \epsilon_2^2$, taking the derivatives of $\psi_2(H)$ and $\psi_3(H)$ with respect to t, we have

$$\left[\frac{\partial \psi_2(H)}{\partial H}\right]^T \frac{\mathrm{d}H}{\mathrm{d}t} = 0, \tag{12}$$

$$\left[\frac{\partial \psi_3(H)}{\partial H}\right]^T \frac{\mathrm{d}H}{\mathrm{d}t} = 0. \tag{13}$$

Substituting (9)–(11) into (12) and (13), we get

$$\begin{aligned}&2ff^T H^T R_c^T R_c\big(-2(H - I + \lambda_1 R_c^T R_c H\\ &\qquad + \lambda_2 R_r^T R_r H) f f^T\big) = 0,\\ &2ff^T H^T R_r^T R_r\big(-2(H - I + \lambda_1 R_c^T R_c H\\ &\qquad + \lambda_2 R_r^T R_r H) f f^T\big) = 0.\end{aligned}$$

It follows that

$$\lambda_1 ff^T H^T (R_c^T R_c)^2 Hff^T + \lambda_2 f(Hf)^T R_c^T R_c R_r^T R_r Hff^T = ff^T H^T R_c^T R_c ff^T - \epsilon_1^2 ff^T, \quad (14)$$

$$\lambda_1 f(Hf)^T R_r^T R_r R_c^T R_c Hff^T + \lambda_2 ff^T H^T (R_r^T R_r)^2 Hff^T = ff^T H^T R_r^T R_r ff^T - \epsilon_2^2 ff^T. \quad (15)$$

Multiplying (14) and (15) by f^T from the right and f from the left, and noting that $f^T f = \|f\|^2$, we have

$$\lambda_1 \|f\|^4 (Hf)^T (R_c^T R_c)^2 Hf) + \lambda_2 \|f\|^4 (Hf)^T R_c^T R_c R_r^T R_r Hf) = \|f\|^4 (Hf)^T R_c^T R_c f - \epsilon_1^2 \|f\|^4, \quad (16)$$

$$\lambda_1 \|f\|^4 (Hf)^T R_r^T R_r R_c^T R_c Hf) + \lambda_2 \|f\|^4 (Hf)^T (R_c^T R_c)^T Hf) = \|f\|^4 (Hf)^T R_r^T R_r f - \epsilon_2^2 \|f\|^4. \quad (17)$$

Defining $\hat{R}_c \overset{\text{def}}{=} R_c^T R_c, \hat{R}_r \overset{\text{def}}{=} R_r^T R_r, \hat{H}_c \overset{\text{def}}{=} H^T R_c^T R_c$ and $\hat{H}_r \overset{\text{def}}{=} H^T R_r^T R_r$, and noting that generally $\|f\| \neq 0$, (16) and (17) are reduced to

$$\|Hf\|^2_{\hat{R}_c^2} \lambda_1 + \|Hf\|^2_{\hat{R}_c \hat{R}_r} \lambda_2 = \|f\|^2_{\hat{H}_c} - \epsilon_1^2, \quad (18)$$

$$\|Hf\|^2_{\hat{R}_r \hat{R}_c} \lambda_1 + \|Hf\|^2_{\hat{R}_r^2} \lambda_2 = \|f\|^2_{\hat{H}_r} - \epsilon_2^2, \quad (19)$$

where $\|f\|^2_X \overset{\text{def}}{=} f^T X f$. The gradient flow dynamic system is (11) with (18) and (19), which is an algebraic differential equation.

From the theory of gradient flows on manifolds (Rapcsak, 1997; Helmke and Moore, 1994), we know that the dynamics of (11) tend to a constant matrix H on the manifold $\|\hat{f}_{ced}\| = \epsilon_1, \ \|\hat{f}_{red}\| = \epsilon_2$. This constant matrix H is the optimal filter in the sense of (5) with (6) and (7).

Theorem 1. *Let an image vector f and two matrices R_c and R_r of proper dimensions be given. If a matrix H on the constraint manifold of (6) and (7) optimizes (5), then H is the solution of the following algebraic differential equation:*

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -2(H - I + \lambda_1 \hat{R}_c H + \lambda_2 \hat{R}_r H) ff^T, \quad (20)$$

$$\|Hf\|^2_{\hat{R}_c^2} \lambda_1 + \|Hf\|^2_{\hat{R}_c \hat{R}_r} \lambda_2 = \|f\|^2_{\hat{H}_c} - \epsilon_1^2, \quad (21)$$

$$\|Hf\|^2_{\hat{R}_r \hat{R}_c} \lambda_1 + \|Hf\|^2_{\hat{R}_r^2} \lambda_2 = \|f\|^2_{\hat{H}_r} - \epsilon_2^2, \quad (22)$$

where $\hat{R}_c = R_c^T R_c, \hat{R}_r = R_r^T R_r, \hat{H}_c = H^T R_c^T R_c$ and $\hat{H}_r = H^T R_r^T R_r$.

The advantage of Theorem 1 is that it converts an optimization problem into that of solving an algebraic differential equation. But, generally, this theorem is more important in a theoretical sense than in real applications. With the high dimensionality of the matrix filter H, solving (20) with (21) and (22) is complex and time consuming.

In applications, we are more interested in the optimal reconstruction of the image vector $\hat{f} \overset{\text{def}}{=} Hf$ than in obtaining the optimal filter H. Let $\hat{f}_e$ denote the block edge vector of $\hat{f}$, which is defined in the same way as the block edge vector f_e of f in Section 2, and let $\hat{f}_{\bar{e}}$ denote the vector whose components are the same as in $\hat{f}$ but not in $\hat{f}_e$. Similarly, let vector $f_{\bar{e}}$ be the vector whose components are the same as in f but not in f_e. From

$$\min \|Hf - f\|^2 = \min \|\hat{f}_e - f_e\|^2 + \min \|\hat{f}_{\bar{e}} - f_{\bar{e}}\|^2,$$

taking account of the constraints (6) and (7), we know that we just need to let $\hat{f}_{\bar{e}} = f_{\bar{e}}$ in the above equation and minimize $\|\hat{f}_e - f_e\|^2$. This means that we only need to minimize $\|\hat{f}_e - f_e\|$ and to let the other components in $\hat{f}$ be equal to their corresponding components in f.

For the vector $\hat{f}_e$, similar to (2) and (3) in Section 2, we have $Q_c \hat{f}_e = \hat{f}_{ced} = R_c \hat{f}$ and $Q_r \hat{f}_e = \hat{f}_{red} = R_r \hat{f}$, where Q_c, Q_r, R_c and R_r are the same matrices as in (2) and (3). So we have the following result:

Theorem 2. *Let an image vector f, and two matrices R_c and R_r of proper dimensions be given. If there is an image vector $\hat{f}$ which satisfies the constraints*

$$\|R_c \hat{f}\| = \epsilon_1, \quad \|R_r \hat{f}\| = \epsilon_2$$

and minimizes $\|\hat{f} - f\|$, then the components of $\hat{f}$ that are not boundary components are equal to the corresponding non-boundary components of f and the block edge vector $\hat{f}_e$ of $\hat{f}$ are determined by the following algebraic differential equation:

$$\frac{\mathrm{d}\hat{f}_e}{\mathrm{d}t} = -2(\hat{f}_e - f_e + \lambda_1 \hat{Q}_c \hat{f}_e + \lambda_2 \hat{Q}_r \hat{f}_e), \quad (23)$$

$$\|\hat{f}_e\|^2_{\hat{Q}_c^2} \lambda_1 + \|\hat{f}_e\|^2_{\hat{Q}_c \hat{Q}_r} \lambda_2 = \hat{f}_e^T \hat{Q}_c f_e - \epsilon_1^2, \quad (24)$$

$$\|\hat{f}_e\|^2_{\hat{Q}_r \hat{Q}_c} \lambda_1 + \|\hat{f}_e\|^2_{\hat{Q}_r^2} \lambda_2 = \hat{f}_e^T \hat{Q}_r f_e - \epsilon_2^2, \quad (25)$$

where $\hat{Q}_c \overset{\text{def}}{=} Q_c^T Q_c, \hat{Q}_r \overset{\text{def}}{=} Q_r^T Q_r$.

Proof. Note that minimizing $\|Hf - f\|^2$ is equivalent to minimizing $\|\hat{f}_e - f_e\|^2$, and constraints $\|R_c \hat{f}\| = \epsilon_1$

and $\|R_r \hat{f}\| = \epsilon_2$ are equivalent to $\|Q_c \hat{f}_e\| = \epsilon_1$ and $\|Q_r \hat{f}_e\| = \epsilon_2$, respectively, as has previously been discussed.

Define

$$\psi_1(\hat{f}_e) = (\hat{f}_e - f_e)^T(\hat{f}_e - f_e),$$

$$\psi_2(\hat{f}_e) = (Q_c \hat{f}_e)^T(Q_c \hat{f}_e),$$

$$\psi_3(\hat{f}_e) = (Q_r \hat{f}_e)^T(Q_r \hat{f}_e).$$

We have

$$\frac{\partial \psi_1(\hat{f}_e)}{\partial \hat{f}_e} = 2(\hat{f}_e - f_e), \tag{26}$$

$$\frac{\partial \psi_2(\hat{f}_e)}{\partial \hat{f}_e} = 2Q_c^T Q_c \hat{f}_e, \tag{27}$$

$$\frac{\partial \psi_3(\hat{f}_e)}{\partial \hat{f}_e} = 2Q_r^T Q_r \hat{f}_e. \tag{28}$$

Let $\psi(\hat{f}_e) = \psi_1(\hat{f}_e) + \lambda_1 \psi_2(\hat{f}_e) + \lambda_2 \psi_3(\hat{f}_e)$, where λ_1 and λ_2 are indeterminates. We have

$$\begin{aligned}\frac{\partial \psi(\hat{f}_e)}{\partial \hat{f}_e} &= \frac{\partial \psi_1(\hat{f}_e)}{\partial \hat{f}_e} + \lambda_1 \frac{\partial \psi_2(\hat{f}_e)}{\partial \hat{f}_e} + \lambda_2 \frac{\partial \psi_3(\hat{f}_e)}{\partial \hat{f}_e} \\ &= 2(\hat{f}_e - f_e + \lambda_1 Q_c^T Q_c \hat{f}_e + \lambda_2 Q_r^T Q_r \hat{f}_e).\end{aligned}$$

So we can compute the gradient flow of $\hat{f}_e$ as

$$\frac{\mathrm{d}\hat{f}_e}{\mathrm{d}t} = -2(\hat{f}_e - f_e + \lambda_1 Q_c^T Q_c \hat{f}_e + \lambda_2 Q_r^T Q_r \hat{f}_e). \tag{29}$$

Next, in the same way as in the proof of Theorem 1, we determine λ_1 and λ_2.

Since $\hat{f}_e$ satisfies $\psi_2(\hat{f}_e) = \epsilon_1^2$ and $\psi_3(\hat{f}_e) = \epsilon_2^2$, we have

$$\left[\frac{\partial \psi_2(\hat{f}_e)}{\partial \hat{f}_e}\right]^T \frac{\mathrm{d}\hat{f}_e}{\mathrm{d}t} = 0, \tag{30}$$

$$\left[\frac{\partial \psi_3(\hat{f}_e)}{\partial \hat{f}_e}\right]^T \frac{\mathrm{d}\hat{f}_e}{\mathrm{d}t} = 0. \tag{31}$$

Substiting (27)–(29) into (30) and (31), we get

$$(Q_c^T Q_c \hat{f}_e)^T(\hat{f}_e - f_e + \lambda_1 Q_c^T Q_c \hat{f}_e + \lambda_2 Q_r^T Q_r \hat{f}_e) = 0,$$

$$(Q_r^T Q_r \hat{f}_e)^T(\hat{f}_e - f_e + \lambda_1 Q_c^T Q_c \hat{f}_e + \lambda_2 Q_r^T Q_r \hat{f}_e) = 0.$$

It follows that

$$\begin{aligned}&\lambda_1 \hat{f}_e^T (Q_c^T Q_c)^2 \hat{f}_e + \lambda_2 \hat{f}_e^T Q_c^T Q_c Q_r^T Q_r \hat{f}_e \\ &\qquad = \hat{f}_e^T Q_c^T Q_c f_e - \|Q_c \hat{f}_e\|^2,\end{aligned} \tag{32}$$

$$\begin{aligned}&\hat{f}_e^T Q_r^T Q_r Q_c^T Q_c \hat{f}_e + \lambda_2 \hat{f}_e^T (Q_r^T Q_r)^2 \hat{f}_e \\ &\qquad = \hat{f}_e^T Q_r^T Q_r f_e - \|Q_r \hat{f}_e\|^2.\end{aligned} \tag{33}$$

Defining $\hat{Q}_c \stackrel{\text{def}}{=} Q_c^T Q_c, \hat{Q}_r \stackrel{\text{def}}{=} Q_r^T Q_r$, we obtain

$$\|\hat{f}_e\|^2_{\hat{Q}_c^2} \lambda_1 + \|\hat{f}_e\|^2_{\hat{Q}_c \hat{Q}_r} \lambda_2 = \hat{f}_e \hat{Q}_c f_e - \epsilon_1^2, \tag{34}$$

$$\|\hat{f}_e\|^2_{\hat{Q}_r \hat{Q}_c} \lambda_1 + \|\hat{f}_e\|^2_{\hat{Q}_r^2} \lambda_2 = \hat{f}_e \hat{Q}_r f_e - \epsilon_2^2. \tag{35}$$

The gradient flow dynamic system is (29) with (34) and (35), which constitutes an algebraic differential equation. ■

The optimal block edge vector $\hat{f}_e$ can substitute the old block edge vector f_e as we reconstruct the decoded image X with the hope of having a higher PSNR.

Since the number of components of $\hat{f}_e$ is much smaller than the number of the entries of the matrix filter H, the computation of $\hat{f}_e$ has much less complexity than the computation in Theorem 1.

4. Algorithm and Experiments

4.1. Algorithm Description

In this section, we propose a numerical algorithm based on (23)–(25). Given a decoded image X with blocking effects, we properly choose two numbers $\epsilon_1 > 0$ and $\epsilon_2 > 0$, an initial vector f_0 with $\|Q_c f_0\| = \epsilon_1$ and $\|Q_r f_0\| = \epsilon_2$, and an appropriate iteration stepsize h first. Then we design the following numerical algorithm with the inputs, $X, f_0, Q_c, Q_r, \epsilon_1, \epsilon_2$ and h, and the output, $\hat{X}$, which is a reconstructed image matrix.

Algorithm:
Compute the block edge vector f_e of the decoded image X. Let $\hat{f}_e = f_0$ as an initial vector. Compute the optimal reconstructed edge vector $\hat{f}_e$ from the following iteration:

Iteration :
do {solveλ_1 and λ_2 first from the subsystem of algebraic equations:

$$\|\hat{f}_e\|^2_{\hat{Q}_c^2} \lambda_1 + \|\hat{f}_e\|^2_{\hat{Q}_c \hat{Q}_r} \lambda_2 = \hat{f}_e^T \hat{Q}_c f_e - \epsilon_1^2,$$

$$\|\hat{f}_e\|^2_{\hat{Q}_r \hat{Q}_c} \lambda_1 + \|\hat{f}_e\|^2_{\hat{Q}_r^2} \lambda_2 = \hat{f}_e^T \hat{Q}_r f_e - \epsilon_2^2,$$

then

$$\hat{f}_e := \hat{f}_e - 2h(\hat{f}_e - f_e + \lambda_1 \hat{Q}_c \hat{f}_e + \lambda_2 \hat{Q}_r \hat{f}_e)$$

}while $\hat{f}_e$ is not a constant vector.

Compute the reconstructed image $\hat{X}$ using $\hat{f}_e$ instead of f_e.

4.2. Experiments

As almost all algorithms reported in the literature for reducing blocking effects are applied to highly compressed images where blocking effects are better evidenced, we also conduct our experiments in highly compressed image cases.

Experiment 1. We use a typical 512×512 Lena image. The image is divided into 8×8 blocks and compressed using the JPEG standard with a compression rate of 32:1. Figure 1 is the 512×512 compressed Lena image. The PSNR of Fig. 1 is 30.3930 dB. Figure 2 is a 128×128 subimage of Fig. 1. It is re-scaled to show the blocking effects. The blocking effects are more clearly visible in Fig. 2. After processing the image by the proposed algorithm with $\epsilon_1 = 8.0503$ and $\epsilon_2 = 5.9053$, the PSNR is 30.9547 dB with an improvement of 0.5617 dB. The improved visual quality can be seen in Fig. 3, which is an 128×128 subimage of the processed 512×512 Lena image, corresponding to the unprocessed image Fig. 2. We also compute the percentage of the processing time for deblocking over the processing time for JPEG decompression. The percentage is 0.16%. It is satisfactory from the viewpoint of applications.

Fig. 1. Image compressed with a rate of 32:1.

Fig. 2. A subimage extracted from Fig. 1.

Fig. 3. Subimage of the image processed in Experiment 1.

Note that since the block edge vector f_e includes information from all block boundaries, the proposed algorithm is to reduce the blocking effects between any adjacent block boundaries simultaneously. We can also define the block edge vector f_e for every adjacent block and reduce them sequentially. That is, for every subimage matrix $(X_{i,j}, X_{i,j+1})$ of the decoded image matrix X in (1) in Section 2, we define $f_e = ((f^r_{i,j})^T, (f^l_{i,j+1})^T)^T$, where $f^r_{i,j}$ and $f^l_{i,j+1}$ are the same as in Section 2, and design a proper matrix Q such that $\|Qf_e\| = \|f^r_{i,j} - f^l_{i,j+1}\| = \epsilon$. We can obtain a new corresponding block edge vector $\hat{f}_e$ between $X_{i,j}$ and $X_{i,j+1}$ using the above algorithm with $Q_c = Q, \epsilon_1 = \epsilon, Q_r$ being the zero matrix and $\epsilon_2 = 0$. Using $\hat{f}_e$ instead of f_e, we can reduce the blocking effects between the block matrices $X_{i,j}$ and $X_{i,j+1}$ in all vertical directions. Similarly, we can reduce the blocking effects sequentially in all horizontal directions. Based on this idea, we design the second experiment with a better PSNR. ⧫

Experiment 2: We use the same Lena image as in Experiment 1 as a test. All the edge difference vector norms of any two adjacent blocks in the original image are calculated and transmitted to the reconstruction end as ϵ. The processed image has an improved PSNR. The PSNR is 31.1339 dB and the improvement value of PSNR is

Fig. 4. Subimage of the image processed in Experiment 2.

0.7401 dB. Figure 4 is an 128×128 subimage of the processed 512×512 Lena image, corresponding to the unprocessed image of Fig. 2. We can see from it the improved reduction of blocking effects. Compared with Fig. 2, Fig. 4 has a much better visual quality. From Fig. 4 we also notice that when the new block edge vector $\hat{f}_e$ replaces the old one, the new blocking artifacts around the new edges are much weaker compared with Fig. 2. The percentage of the processing time for deblocking over the processing time for JPEG decompression is 0.24%. Compared with the corresponding number 0.16% in Experiment 1, 0.24% is a little greater. But if we compare the two experiments with respect to the PSNR, Experiment 2 is better than Experiment 1. So in applications, we must attain some tradeoff between the processing time and the image quality.

5. Conclusion

In this paper, a gradient flow optimization method was applied to the problem of reducing blocking effects in transform coding. By using properties of gradient flows, an optimal filter design method and an optimal reconstruction method for reducing the blocking effects were presented. These design methods were based on dynamic algebraic differential equations, which can be turned into iterative formulas when they are used in numerical computations. An algorithm for the optimal reconstruction of images was provided. This algorithm can be used to reconstruct a decoded image by using only two constrained parameters. Experiments showed that the algorithm can obtain significant improvements regarding the PSNR and good visual quality.

References

Helmke U. and Moore J.B. (1994): *Optimization and Dynamical Systems*. — London: Springer-Verlag.

ISO (1991) : Commitee Draft ISO/IEC CD 10918-1, *Digital compression and coding of continuous-tone still images, Part 1: Requirements and guidelines,* March 15, 1991.

ISO (1993): ISO/IEC JTC1/SC29/WG11, *Test Models 5, MPEG 93/457, Document AVC-491,* April,1993.

Jain A.K. (1989): *Fundamentals of Digital Image Processing*. — Englewood Cliffs, NJ: Prentice-Hall.

Kim T.K., Paik J.K., Won C.S., Choe Y.S., Jeong J. and Nam J.Y. (2000): *Blocking effect reduction of compressed images using classification-based constrained optimization*. — Sign. Process. Image Comm., Vol. 15, pp. 869–877.

Minami S. and Zakhor A. (1995): *An optimization approach for removing blocking effects in transform coding*. — IEEE Trans. Circ. Syst. [Video Technol]., Vol. 5, No. 2, pp. 74–82.

Rao K.K. and Hwang J.J. (1996), *Techniques and Standards for Image, Video and Audio Coding*. — Englewood Cliffs, NJ: Prentice Hall, Inc.

Rapcsak T. (1997), *Smooth Nonlinear Optimization in* $\mathbb{R}^n$. — Dordrecht: Kluwer.

Reeve H.C. and Lim J.S. (1984): *Reduction of blocking effects in image coding*. — Opt. Eng., Vol. 23, No. 1, pp. 34–37.

Rosenholtz R. and Zakhor A. (1992): *Iterative procedures for reduction of blocking effects in transform image coding*. — IEEE Trans. Circ. Syst. Video Technol., Vol. 2, No. 1, pp. 91–94.

Won C.S. and Derin H. (1992): *Unsupervised segmentation of noisy and textured images using Markov random fields*. — CVGIP: Graph. Mod. Image Process., Vol. 54, No. 4, pp. 308–328.

Yang Y., Galatsanos N.P. and Katsaggelos A.K. (1993): *Regularized reconstruction to reduce blocking artifacts of block discrete cosine transform compressed images*. — IEEE Trans. Circ. Syst. [Video Technol]., Vol. 3, No. 6, pp. 421–432.

Yang Y., Galatsanos N.P. and Katsaggelos A.K. (1995): *Projection-based spatially adaptive reconstruction of block-transform compressed images*. — IEEE Trans. Image Process., Vol. 4, No. 7, pp. 896–908.

Received: 7 February 2003
Revised: 29 December 2003