Maximum-Likelihood Parameter Estimation for Image Ringing-Artifact Removal

Seungjoon Yang, Yu-Hen Hu, *Fellow, IEEE*, Truong Q. Nguyen, *Senior Member, IEEE*, and Damon L. Tull, *Member, IEEE*

***Abstract*—At low bit rates, image compression codecs based on overlapping transforms introduce spurious oscillations known as *ringing artifacts* in the vicinity of major edges. Unlike previous works, we present a maximum-likelihood approach to the ringing-artifact removal problem. Our approach employs a parameter-estimation method based on the k-means algorithm with the number of clusters determined by a cluster-separation measure. The proposed algorithm and its simplified approximation are applied to JPEG2000 compressed images. Our results show effective and efficient removal of ringing artifacts.**

***Index Terms*—Hierarchical clustering, image ringing-artifact, maximum-likelihood estimation, robust filter.**

I. Introduction

JPEG introduces a blocking artifact at medium and low bit-rates because of its short and nonoverlapping basis functions. Overlapping transforms such as wavelet and GenLOT [1] are introduced to reduce or eliminate the blocking artifact, but they have spurious oscillations in the vicinity of major edges at low bit-rates. Such a coding artifact is called the *ringing artifact*, which is due to the abrupt truncation of the high frequency wavelet coefficients.

In practice, it is desirable to obtain an image free of compression related artifacts, thus improving image quality at low bit rates. An artifact free image can be estimated from the compressed image by maximum *a posteriori* (MAP) estimation [2]–[4]. The problem is to estimate the artifact free image f given a compressed image g. In MAP estimation approaches, the quantity being estimated, f, is considered to be a random variable whose property is modeled by a probability density, following the Bayesian viewpoint. The MAP estimate is the f that maximizes the posterior probability that is expressed in terms of the conditional probability and the prior probability. For transform based codecs, the conditional probability is modeled in the transform domain, while the prior probability is modeled in the spatial domain. This aspect significantly increases the computational complexity when the solution is obtained through iterative algorithms.

In this work, we employ a maximum likelihood (ML) estimation viewpoint where the quantity being estimated, f in our problems, is regarded as a fixed but unknown quantity. This shift in viewpoint requires a new formulation to describe the degradation process. We assume that the image is a montage of flat surfaces. The flat surface model is fitted to the observed image by estimating the parameters of the model. We employ the k-means algorithm [5] in conjunction with a hierarchical clustering algorithm to achieve this fit. A cluster separation measure (CSM) [6], modified for elimination of ringing artifact, is used in the hierarchical clustering algorithm to determine the number of flat surfaces. The model parameters are estimated by the k-means algorithm. The proposed method is simplified by using a three cluster model whose cluster centers are initialized by a deterministic rule. The simplified version of the proposed method is a noniterative mapping between the neighboring pixel values and the estimate. It is particularly attractive for its good performance and simplicity. The proposed methods are applied to remove the ringing artifact in images compressed by JPEG2000 [7]. The results show effective and efficient elimination of ringing artifact with improved image quality.

Section II reviews existing methods of coding-artifact removal, including MAP estimation approaches and projection onto convex sets (POCS). Section III introduces a ML viewpoint to the artifact removal problem and presents the k-means algorithm and the hierarchical clustering algorithm required in our approach. A simplified ML algorithm based on a three cluster model is presented. This algorithm is termed the robust filter. Section IV includes the result of the hierarchical clustering algorithm and the results of ringing-artifact removal on JPEG2000 compressed images. Section V concludes the paper.

II. Review of Existing Methods

Transform based image compression codecs can be modeled by

$$g = T^{-1}QTf \tag{1}$$

where

- g compressed image;
- f original image;
- T forward transform;
- T^{-1} inverse transform;
- Q quantization operation.

Coding artifacts introduced by image compression codecs can be removed by MAP estimation [2], [4] or POCS [8], [9], where the estimate of the original image before the compression is obtained given a compressed image and the codec model. Since the quantization operation is a many-to-one mapping, the

Manuscript received March 28, 2000; revised March 27, 2001. This paper was recommended by Associate Editor A. Kot.

S. Yang, Y.-H. Hu, and D. L. Tull are with the Department of Electrical and Computer Engineering, University of Wisconsin–Madison, Madison, WI 53706 USA (e-mail: hu@engr.wisc.edu).

T. Q. Nguyen is with the Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215 USA.

Publisher Item Identifier S 1051-8215(01)06528-4.

estimation problem is ill-posed. By using prior information on the estimate, the ill-posed problem is avoided.

A. MAP Estimation

In MAP estimation approaches, the problem of estimating the original image $\hat{f}$, given a compressed image g, is written as [2]

$$\hat{f} = \arg\max_f \log P[f|g] \tag{2}$$

$$= \arg\max_f \{\log P[g|f] + \log P[f]\}. \tag{3}$$

The conditional probability $P[g|f]$ depends on the specifics of the codec. For a transform based image compression codec modeled by (1), the conditional probability $P[g|f]$ is

$$P[g|f] = \begin{cases} \text{constant}, & f \in \Omega \\ 0, & f \notin \Omega \end{cases} \tag{4}$$

where

$$\Omega = \{f|g = T^{-1}QTf\}. \tag{5}$$

Under this conditional probability $P[g|f]$, all the images in the feasible set Ω in (5) are equally probable. Regularization allows us to obtain a unique, meaningful solution out of all the images in the feasible set.

In MAP estimation approaches, the quantity f that we are estimating is considered to be a random variable, and the prior knowledge is described by the probability density $P[f]$. If the probability $P[f]$ follows the Gibbs distribution [10], then

$$P[f] \propto \exp\left\{-\sum_{c\in\mathcal{C}} \rho(d_c f)\right\} \tag{6}$$

where

- $\mathcal{C}$ set of cliques that belong to a neighborhood system;
- ρ clique potential function;
- d_c differential operator that models the dependency of the pixels in the clique.

The operator d_c is usually chosen such that the prior distribution models the smoothness of the image. Robust forms of potential functions ρ [11], [12] are used to impose weak continuity on the solution and to allow for edges in the MAP estimate.

Using the conditional probability $P[g|f]$ in (4) and the prior probability $P[f]$ in (6), the MAP estimation problem in (3) becomes

$$\hat{f} = \arg\min_{f\in\Omega} \sum_{c\in\mathcal{C}} \rho(d_c f). \tag{7}$$

In [2], the Huber function is used as the potential function. In [4], a set of functions derived from the line process is used adaptively according to the local statistics of the image.

B. POCS

In POCS approaches [8], [9], the problem of obtaining an artifact-free image is to find an image that satisfies the following constraints:

$$T^{-1}QTF = g$$
$$||Cf|| \le \varepsilon \tag{8}$$

where C is derived to explicitly model the severity of the degradation coming from the coding artifact. The sets that define the above constraints are given as

$$\Omega_d = \{f|T^{-1}QTf = g\} \tag{9}$$

$$\Omega_s = \{f|||Cf|| \le \varepsilon\} \tag{10}$$

where the recovered image f is constrained to be in the intersection of Ω_d and Ω_s.

The set Ω_d is the set of feasible images, the set of all the images that should compress to the observed compressed image g. Knowledge regarding the estimated image is imposed by the set Ω_s. In previous work, for example [8], C is explicitly designed to counter blocking artifacts. This is possible since block locations are known a priori in block based codecs. This is not the case in codecs based on overlapping transforms. The location of the distortion is not fixed by the structure of the bit stream.

Contrary to the MAP estimation approaches, the quantity f being estimated is not a random variable that is described by the probability density. In POCS, f is a vector with an unknown value whose properties are described by the sets Ω_d and Ω_s.

C. Motivation for a New Approach

Both MAP estimation and POCS approaches require iterative algorithms in finding a solution. They need extra loads of computation in addition to the usual loads of iterative algorithms. The algorithms operate on two different domains. The enforcement of the feasible image set Ω given in (5) and the set Ω_d given in (9) operates on the transform domain. On the other hand, the enforcement of our prior knowledge of smoothness of image is enforced in the spatial domain by (6) for the MAP estimation and by (10) for POCS. Hence, each iteration involves the addition of forward and inverse transforms to switch back and forth between two domains. This aspect increases the computational complexity significantly.

III. Proposed Method

In MAP estimation approaches, the quantity f that we are estimating is considered to be a random variable whose property is described by the probability density $P[f]$. This view point is based on the Bayesian estimation. Another approach to the problem is the ML estimation, where the quantity being estimated is viewed as a fixed but unknown value. ML estimation is usually simpler than Bayesian estimation. This paper adopts the ML estimation viewpoint, and the image f is considered to be an unknown quantity.

In many instances, the purpose of ringing-artifact removal is to replace a rippled surface with a flat one. We assume that the estimate f is a montage of flat surfaces. In order to manage a broad class of images, a flat surface model is applied locally to small regions of the image.

We begin with a window of size $[w \times w]$ centered at (i, j)th pixel. This window slides through the compressed image g pixel by pixel to pick samples $\mathcal{G}$. A flat surface model consists of the number of surfaces K, grayscale values of each surface θ, and corresponding surface information s. The grayscale values of each surface form a $[K \times 1]$ vector θ. The surface information s is a $[w \times w]$ size matrix with its elements taking the values in

$\{1,\ldots,K\}$. The flat surface model image of size $[w \times w]$ can be written as

$$S = \sum_{k=1}^{K} \theta_k \mathbf{1}_{s=k} \tag{11}$$

where $\mathbf{1}$ is a vector valued indicator function. The computation of $\mathcal{S}$ from (11) requires an estimate of s and θ. After the flat surface model image $\mathcal{S}$ is obtained, the center pixel of $\mathcal{S}$, denoted by $\mathcal{S}_c$, is taken as the (i,j)th pixel of the ringing-artifact-free image $\hat{f}$.

The samples $\mathcal{G}$ taken from the compressed image are regarded as incomplete data with the cluster information s missing. We consider a problem of estimating the parameter θ from the complete data $(\mathcal{G}, z)$. The ML estimation of the model parameter θ can be written as

$$\hat{\theta} = \arg\max_{\theta} P\left[\mathcal{G}|\theta\right]. \tag{12}$$

By adding the missing data to the problem, it becomes

$$\hat{\theta} = \arg\max_{\theta} P\left[\mathcal{G}, s|\theta\right]. \tag{13}$$

The order of the flat surface model, i.e., the number of surface K, is another model parameter that has to be determined. First, we introduce a parameter estimation method assuming the order of the model is given. The determination of the order of the model will be discussed in Section III-B.

A. k-Means Algorithm

One way to solve the problem with missing data is the expectation maximization (EM) algorithm that consists of the following two steps:

$$\text{E-Step: } \mathcal{Q}\left(\theta|\hat{\theta}^{\tau}\right) = E\left[\log P\left[\mathcal{G}, s|\theta\right]|\mathcal{G}, \hat{\theta}^{\tau}\right] \tag{14}$$

$$\text{M-Step: } \hat{\theta}^{\tau+1} = \arg\max_{\theta} \mathcal{Q}\left(\theta|\hat{\theta}^{\tau}\right). \tag{15}$$

In E-Step of the EM algorithm, we take expectation of the log-likelihood function with respect to the missing data s. There is a probability related to the surface information s. The decision on the surface configuration is through the probability, where a particular configuration of s is given with a corresponding probability. Hence, it is a soft decision on the surface information. In practice, taking the expectation over all possible surface information s is not applicable because of the huge number of possible surface configurations. If we replace the soft decision with the hard decision based on which configuration is most probable, there is no need for taking the expectation. Since the expected value is going to be dominated by the most probable surface configuration, the replacement of the decision provides a good approximation. This approximation of the EM algorithm is called the k-means algorithm [5]. The approximation of the EM steps in the k-means algorithm are

$$\text{E-Step: } \hat{s}^{\tau+1} = \arg\max_{z} \log P\left[\mathcal{G}, s|\hat{\theta}^{\tau}\right] \tag{16}$$

$$\text{M-Step: } \hat{\theta}^{\tau+1} = \arg\max_{\theta} \log P\left[\mathcal{G}, \hat{s}^{\tau+1}|\theta\right]. \tag{17}$$

Note that, in the E-Step of the k-means algorithm, we determine the most probable surface configuration. Clustering algorithms based on the k-means algorithm have been used in image segmentation [13]–[16]. We adopt the k-means algorithm for its simplicity and good approximation of the EM algorithm [14].

B. Cluster Validation

This section addresses the determination of the order of the flat surface model. Removal of the ringing artifact depends on the order of the model. For example, Fig. 1 shows degradation by the ringing artifact on a synthetic image. In order to restore a ringing-artifact-free image, the number of surfaces in the flat surface model has to be chosen correctly as two. As the window slides through the image to pick up the samples $\mathcal{G}$, the statistics of samples $\mathcal{G}$ may change. Hence, the number of surfaces may also change and must be determined from the samples $\mathcal{G}$.

The cluster validation is the problem of finding the number of clusters in cluster analysis, or the number of surfaces in our problem. Partitional clustering methods require input from the user to determine the number of clusters. Hierarchical clustering methods produce nested partitions of one to $K_{\max}$ and determine the number of clusters through a criterion of merit [17].

1) A Hierarchical K-Cluster Model: One of the criteria of merit is to use the cluster separation measure (CSM) [6]. Similarity between two clusters is defined by

$$R_{k,l} = \frac{\sigma_k + \sigma_l}{|\theta_k - \theta_l|} \tag{18}$$

where σ^2 is the variance of the samples in the cluster. $R_{k,l}$ is an intuitive measure of similarity. When the distance between the clusters increases while the variances remain the same, similarity decreases. When the variances decrease while the distance remains the same, similarity decreases. For a given number of clusters $L \in \{2,\ldots,K_{\max}\}$, the average CSM is

$$\text{CSM}(L) = \frac{1}{L}\sum_{k=1}^{L} \max_{\substack{l\in\{1,\ldots,K_{\max}\}\\ l\neq k}} R_{k,l}. \tag{19}$$

The number of the cluster is determined by

$$K = \min_{L\in\{2,\ldots,K_{\max}\}} \text{CSM}(L). \tag{20}$$

CSM is suitable for our problem of ringing-artifact removal because it uses the distance between the clusters. A ringing-artifact-free image can be obtained by prohibiting surfaces with similar grayscales. Hence, we should prohibit surfaces with centers that are near to each other. To achieve this, the similarity in (18) is modified to

$$R_{k,l} = \frac{\sigma_k + \sigma_l}{\varrho(\theta_k - \theta_l)}. \tag{21}$$

The function ϱ is given by

$$\varrho(x) = \begin{cases} c, & \text{if } |x| \le \gamma \\ |x|, & \text{otherwise} \end{cases} \tag{22}$$

where γ is a positive parameter and c is a small positive constant less than γ. The use of the function ϱ makes the clusters whose center distances are less than γ equally similar to each other.

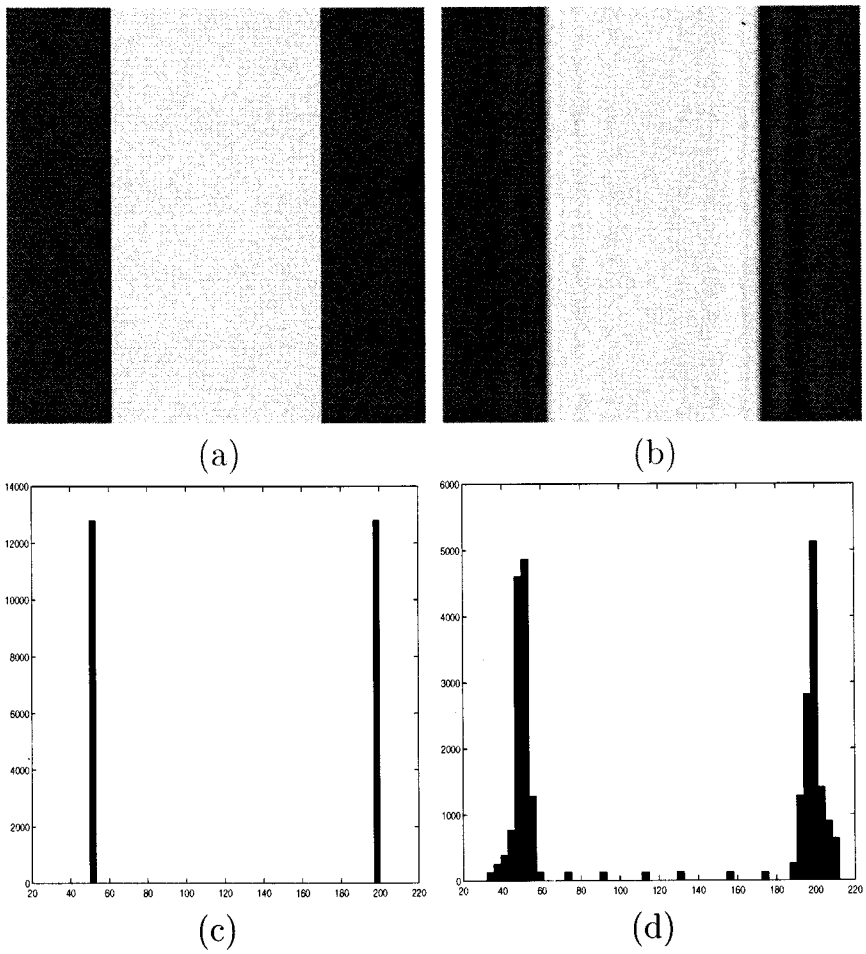

Fig. 1. Degradation by the ringing artifact, an example with a synthetic test images. (a) Original image. (b) Compressed image. (c) Histogram of original image. (d) Histogram of compressed image.

With c being a small number, the number of clusters L with clusters close together is less likely to be chosen as the order of the flat surface model K. The function is a variant of Vapnik's ϵ-insensitive function [18].

The first way to solve the problem in (13) is to use the k-means algorithm, with the number of surfaces determined by the hierarchical clustering method. The algorithm is summarized by the following.

STEP 1:

Run k-means algorithm with L clusters.

Measure the maximum similarity between every two cluster pairs.

$$R[k] = \max_{\substack{l=\{1,\ldots,L\} \\ l \neq k}} R_{k,l}, \quad \text{for } k = 1, \ldots, L$$

Take the average.

$$\mathrm{CSM}[L] = \mathrm{ave}_{k=\{1,\ldots,L\}} R[k]$$

Merge two clusters with largest $R[k]$.

}

STEP 2: Determine the number of the cluster.

$$K = \min_{L=\{2,\ldots,K_{\max}\}} \mathrm{CSM}(L)$$

STEP 3: Run k-means algorithm with K clusters.

The limitation of CSM is that it requires at least two clusters. In order to make a one cluster model available, we add a decision step in STEP 2. If the minimum occurs at $K = 2$ and the distance between the centers of two clusters are smaller than γ, the two clusters are merged to become one cluster. Another limitation of the CSM is that it requires at least one sample in each cluster. Following the method in [6], arbitrary large variance is applied to the cluster with one sample. Since the samples $\mathcal{G}$ are from the compressed image, there are cases when all the samples in a cluster have the same grayscale values making its variance zero. One cannot measure the similarity between the clusters when both have zero variances. To avoid this case, variances of those clusters are set to one.

2) Three-Cluster Model: A Robust Filter: The second algorithm to solve the problem in (13) is developed by using a three-cluster model with the cluster centers determined by a

simple rule. Given the samples $\mathcal{G}$, the cluster centers are initialized as

$$\theta = \begin{bmatrix} \mathcal{G}_c - 2_\gamma \\ \mathcal{G}_c \\ \mathcal{G}_c + 2_\gamma \end{bmatrix} \tag{23}$$

where $\mathcal{G}_c$ denotes the grayscale value of the center pixel in the window. Furthermore, the number of iteration in the k-means algorithm is set to one. The estimate is still an ML estimate under the probability density $P[\mathcal{G}|\theta]$ approximated by the abbreviated k-means algorithm.

Note that the center pixel of the window $\mathcal{S}_c$ is taken as the (i,j)th pixel of the ringing-artifact free image $\hat{f}$. Therefore, θ which $\mathcal{S}_c$ takes is the only parameter of interest. In Section III-B-1, a hierarchical algorithm is used in order to single out the samples that belong to the same cluster as $\mathcal{S}_c$. In the three cluster model, we assume that the samples in $\mathcal{G}$ that are closer to the center pixel $\mathcal{G}_c$ in grayscale values than γ belongs to the same cluster. Then the parameter of our interest becomes a mean of the samples in that cluster. A mapping between the samples in $\mathcal{G}$ and the value $\mathcal{S}_c$, or the (i,j)th pixel of $\hat{f}$, is called the robust filter. The robustness is in that the effect of the samples outside of that cluster on the estimate is limited or reduced.

The operation of the robust filter can be written explicitly. Let $\mathcal{C}(i,j)$ denote the index set of pixels in $\mathcal{G}$ centered at the (i,j)th pixel. Define the index set $\mathcal{I}(i,j;\gamma)$ such that

$$\mathcal{I}(i,j;\gamma) = \{(p,q)|(p,q) \in C(i,j), |g_{p,q} - g_{i,j}| \leq \gamma\}. \tag{24}$$

For the entire image, the robust filter can be written as

$$\hat{f} = \arg\min_f \sum_{i,j} \sum_{m,n \in \mathcal{C}(i,j)} (g_{m,n} - f_{i,j})^2 \mathbf{1}_{m,n \in \mathcal{I}(i,j;\gamma)} \tag{25}$$

where $\mathbf{1}$ is the indicator function and $\sum_{i,j}$ is over the entire image. Denote the function V_c by

$$V_c(g;f) = \sum_{i,j} \sum_{m,n \in \mathcal{C}(i,j)} (g_{m,n} - f_{i,j})^2 \mathbf{1}_{m,n \in \mathcal{I}(i,j;\gamma)}. \tag{26}$$

We regard f as a parameter of the function V_c, and adopt a notation $V_c(g;f)$ instead of $V_c(g,f)$. The robust filter is equivalent to the ML estimation of the image with the probability $P[g|f]$ modeled by the potential function $V_c(g;f)$, such that

$$P[g|f] \propto \exp\{-V_c(g;f)\}. \tag{27}$$

The estimate is

$$\hat{f} = \bar{f}^\gamma \tag{28}$$

where $\bar{f}^\gamma$ is the conditional mean defined by

$$\bar{f}^\gamma_{i,j} = \frac{\sum_{p,q \in \mathcal{I}(i,j;\gamma)} g_{p,q}}{\#_{\mathcal{I}(i,j;\gamma)}} \tag{29}$$

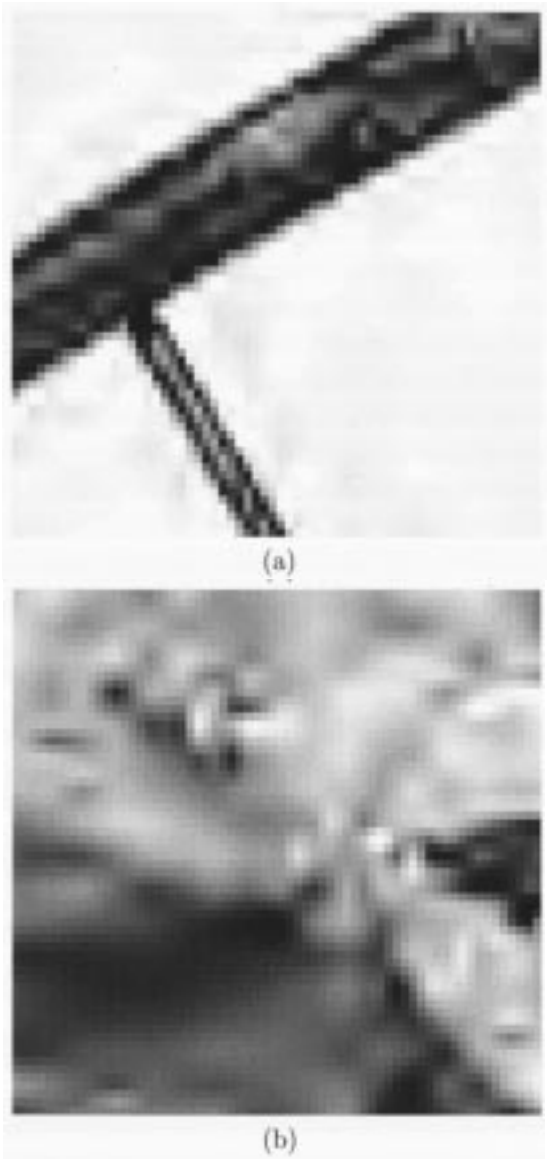

Fig. 2. Test images for hierarchical clustering algorithm, parts of the *bike* image. Ringing artifact: (a) visible oscillation near edges and (b) masked by texture—not objectionable.

where $\#_{\mathcal{I}(i,j;\gamma)}$ is the number of pixels in the set $\mathcal{I}(i,j;\gamma)$.

The estimate is just a conditional mean of neighboring pixels. Its computational complexity is far less than that of the algorithm with K cluster model, MAP estimation-based algorithms, or POCS-based algorithms.

IV. Experiments and Results

In this section, we demonstrate the validity of the proposed methods. Section IV-A explains how the hierarchical clustering algorithm determines the order of the flat surface model with examples. Section IV-B shows the results of the ringing-artifact removal algorithms based on both the hierarchical clustering and the three cluster model with JPEG2000 compressed images. In all the experiments, we employ a Gaussian mixture model for $P[\mathcal{G}, s|\theta]$.

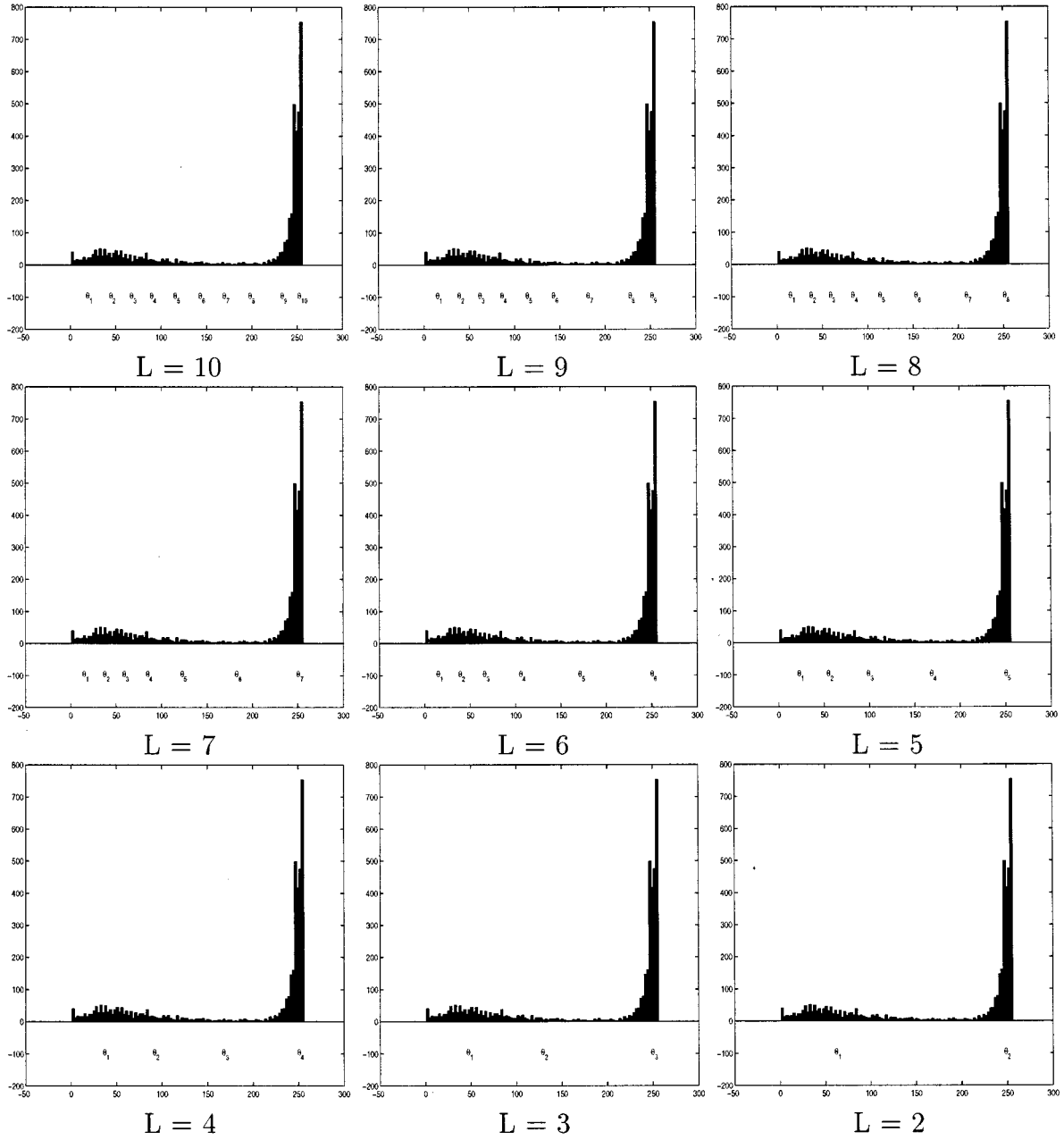

Fig. 3. Example of hierarchical clustering algorithm with the test image (a), cluster center θ for $L = \{2, \ldots, K_{\max}\}$ are shown at the bottom of histograms. $K_{\max} = 10$.

A. Hierarchical Clustering

Fig. 2(a) is a part of the *bike* image, which is compressed by JPEG2000 at 0.125 bits per pixel (bpp). This part of the image shows visible oscillation near edges due to the ringing artifact. The hierarchical clustering algorithm is initiated with $L = K_{\max}$ and equally spaced cluster centers. $K_{\max}$ should be larger than the actual model order. Large $K_{\max}$ also alleviates the problem involved with the initialization of the cluster centers in the k-means algorithm. In the experiments, $K_{\max}$ is set to 10. The algorithm finds the cluster center θ with the given L cluster model, and measures the average CSM. The two clusters which are most similar to each other are merged, and the algorithm continues with the $L-1$ cluster model until $L = 2$. Fig. 3 shows

TABLE I
EXAMPLE OF HIERARCHICAL CLUSTERING ALGORITHM WITH THE TEST IMAGE (A), AVERAGE CSM, THE MINIMUM OCCURS AT $L = 2$. THEREFORE, THE ORDER OF FLAT SURFACE MODEL K IS CHOSEN AS TWO

L	average CSM
10	∞
9	1.0690
8	0.9514
7	0.8247
6	0.7129
5	0.4175
4	0.3943
3	0.4865
2	0.2726

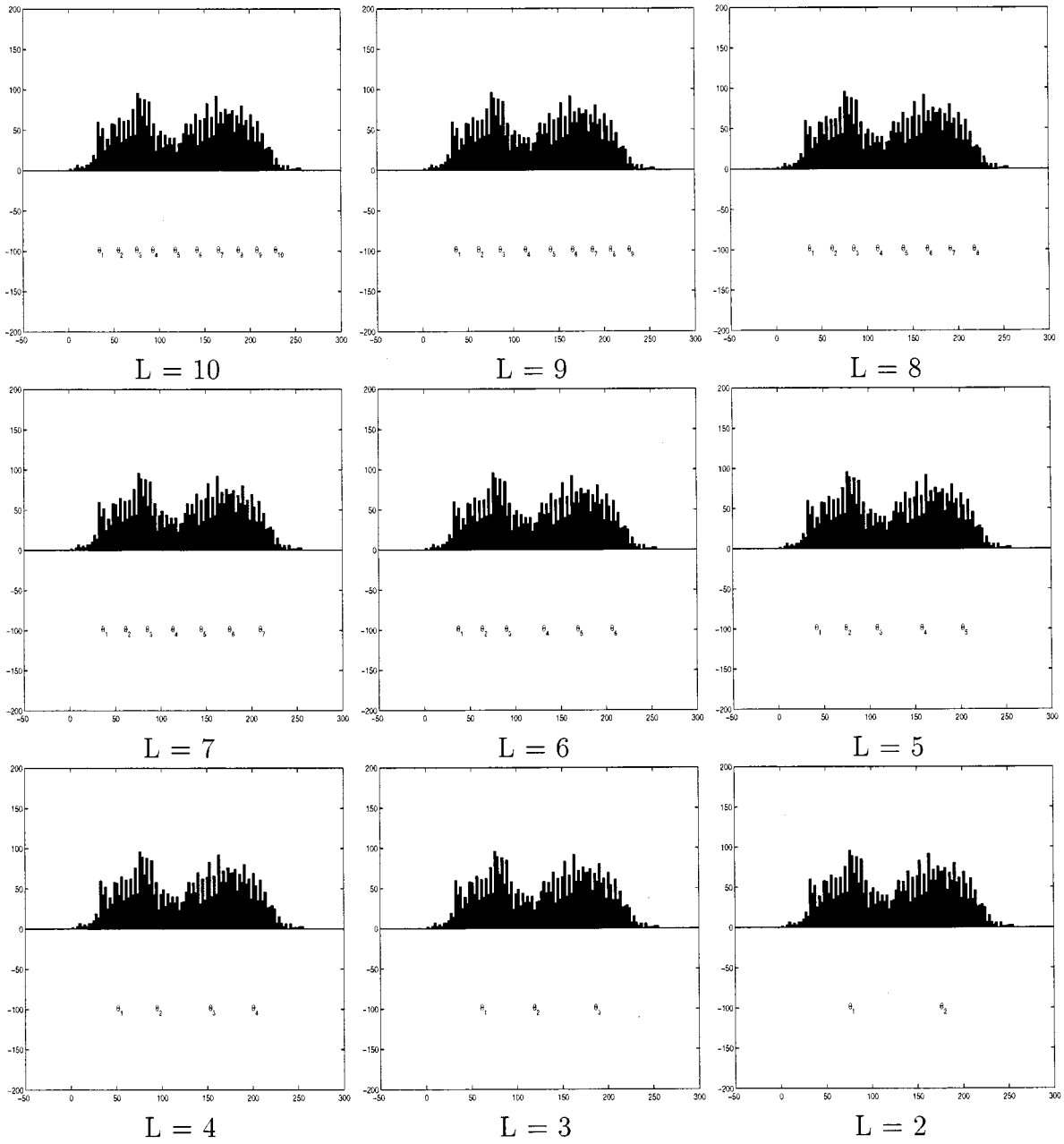

Fig. 4. Example of hierarchical clustering algorithm with the test image (b), cluster center θ for $L = \{2, \ldots, K_{\max}\}$ are shown at the bottom of histograms ($K_{\max} = 10$).

the histograms and the cluster centers with $L = \{2, \ldots, K_{\max}\}$. Table I shows the average CSM for each cluster model. For the $L = 10$ case, there are clusters whose center distance is less than γ. With c chosen as a very small number, the average CSM tends to infinity. The minimum of the average CSM occurs with the two cluster model. The hierarchical clustering algorithm picks the two cluster model for the given image. Hence, $K = 2$. With the two cluster model, the ringing around the edge will be removed.

Fig. 2(b) is another part of the *bike* image at 0.125 bpp. This part of the image contains texture. The ringing artifact in this part of the image, if there is any, is not objectionable because it is masked by the texture. This part of image should be modeled with a sufficiently large number of clusters so that texture is preserved. The same experiment is repeated to demonstrate the validity of the hierarchical clustering algorithm. Fig. 4 shows the histograms and the cluster centers, and Table II shows the average CSM. The minimum of the average CSM occurs at $L = 8$. Hence, the order of the model K is set at eight.

Two examples, one for the region containing major edge and the other for the region containing texture, show that the hierarchical clustering algorithm operates correctly and can be used to estimate the order of our flat surface model.

B. Image Ringing-Artifact Removal

The proposed methods are applied to the ringing-artifact removal of JPEG2000 compressed images. A set of test images consists of the *Aerial2, Bike, Café, Chart, Target, Tools,* and

TABLE II
EXAMPLE OF HIERARCHICAL CLUSTERING ALGORITHM WITH THE TEST IMAGE (B), AVERAGE CSM, THE MINIMUM OCCURS AT $L = 8$. THEREFORE, THE ORDER OF FLAT SURFACE MODEL K IS CHOSEN AS EIGHT

L	average CSM
10	∞
9	0.6249
8	0.4586
7	0.5281
6	0.8312
5	0.4588
4	0.4791
3	0.6240
2	0.5375

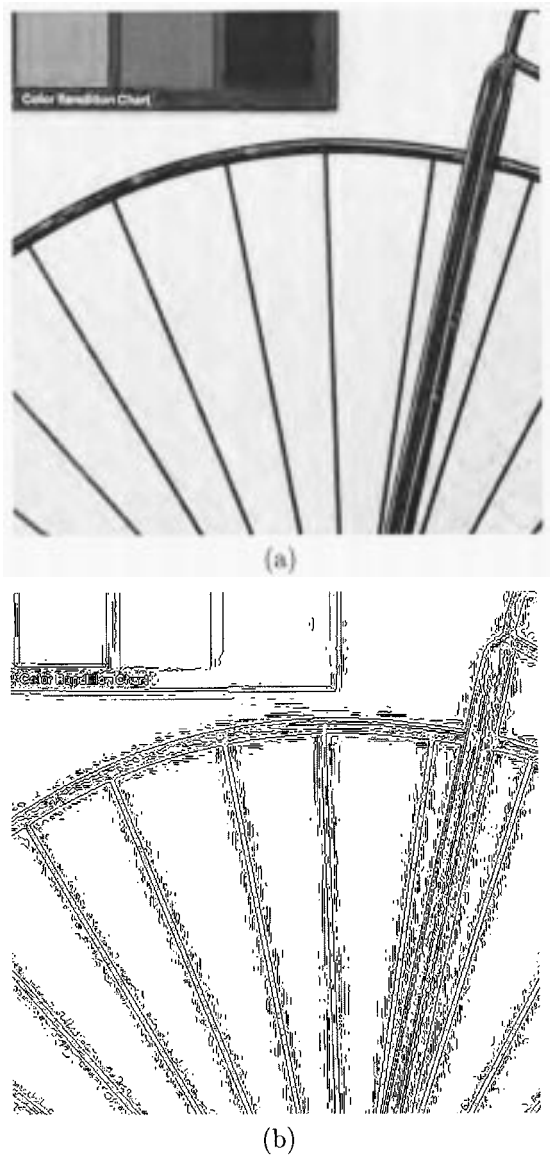

Fig. 5. Compressed image, a part of the *bike* image at 0.125 bpp. (a) Image. (b) Edge map.

Woman images. Images are compressed at 0.125 and 0.0625 bpp, and the proposed methods are applied to remove the ringing artifact.

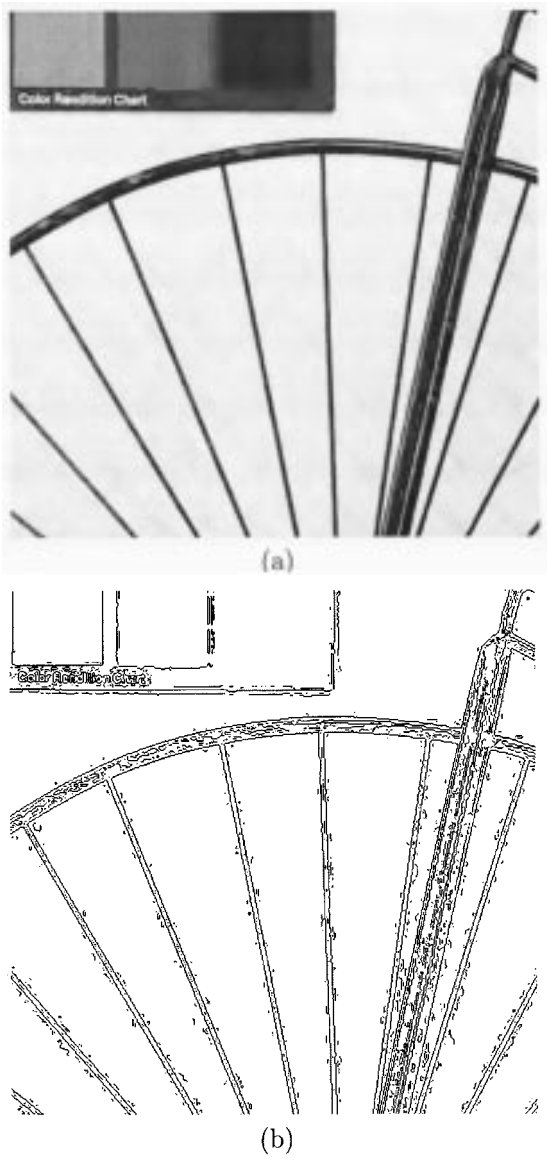

Fig. 6. Image post-processed by the k-means algorithm and the hierarchical clustering based method, a part of the *bike* image at 0.125 bpp. (a) Image. (b) Edge map.

Fig. 5 is a part of the *bike* image compressed at 0.125 bpp. The image in Fig. 5(a) shows the ringing around the edges, and the edge map in Fig. 5(b) shows the false edges introduced by the ringing artifact. Fig. 6 shows the result of the proposed method with the hierarchical K cluster model. Fig. 7 shows the result of the robust filter. Images in Fig. 7(a) in both figures show effective removal of the ringing artifact, which is validated by the removal of false edges in the edge maps shown in Fig. 7(b). Subjective image quality is improved greatly after the elimination of the ringing artifact.

The visible ringing measure (VRM) [19] is an objective distortion measure for ringing artifacts. It is the average local variance calculated on a small-size window in the vicinity of major edges. Ringing-artifact-prone regions are detected by a series of morphological operations. First, an edge map is generated by the sobel operator. After the edge map is cleaned by de-noising and line–curve linking, the edge map is dilated to indicate the

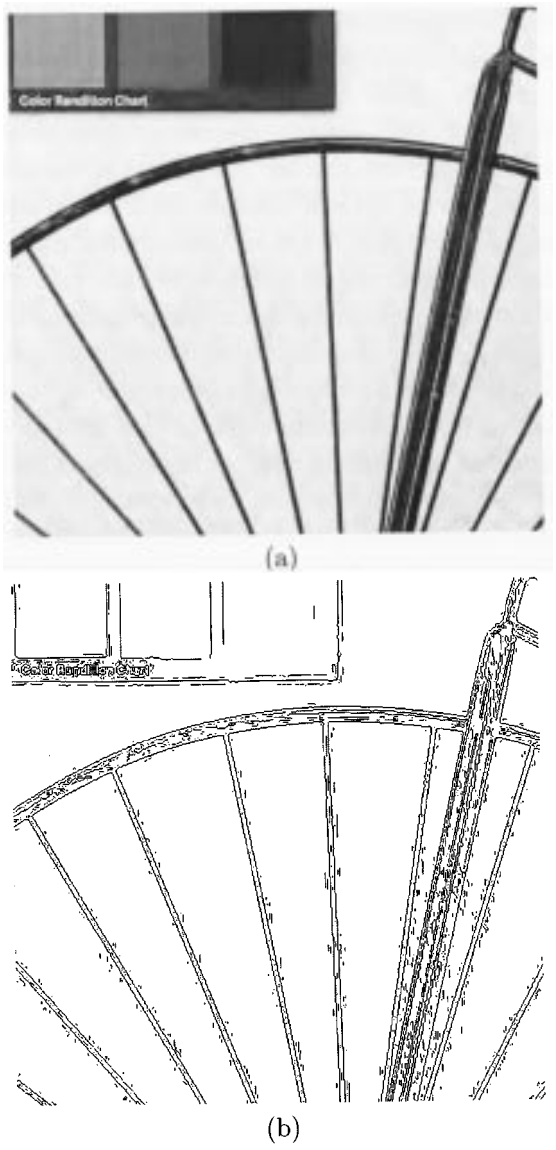

Fig. 7. Image post-processed by the robust filter, a part of the *bike* image at 0.125 bpp: (a) image and (b) edge map.

region around the major edges. Then the actual edges are excluded. Variations of grayscale values inside the detected region are due to the oscillations introduced by the ringing artifacts. Therefore, the average variance of the pixel values in this region can be used to quantify the degradation. Each step involved in VRM is shown in Fig. 8. In [19], a sequence of morphological operations for removal of the ringing artifact is also provided. Tables III and IV show a decrease in VRM after post-processing the images by the proposed methods, which confirms the removal of the ringing artifact.

One of the major advantages of the robust filter is its simplicity. Its computational complexity is far less than all POCS or MAP-based post-processing schemes. One of the measures of complexity is processing time compared to decoding time. Processing time of the robust filter is about the same as decoding time. The results are shown in Table V. Both decoding time and processing time include file i/o. The program is written in straight forward C++ and tested on a 333-MHz dual-Pentium PC with 512 M RAM running on Windows NT 4.0.

Another issue in terms of implementation is memory usage. Most of the iterative algorithms need an image size buffer to

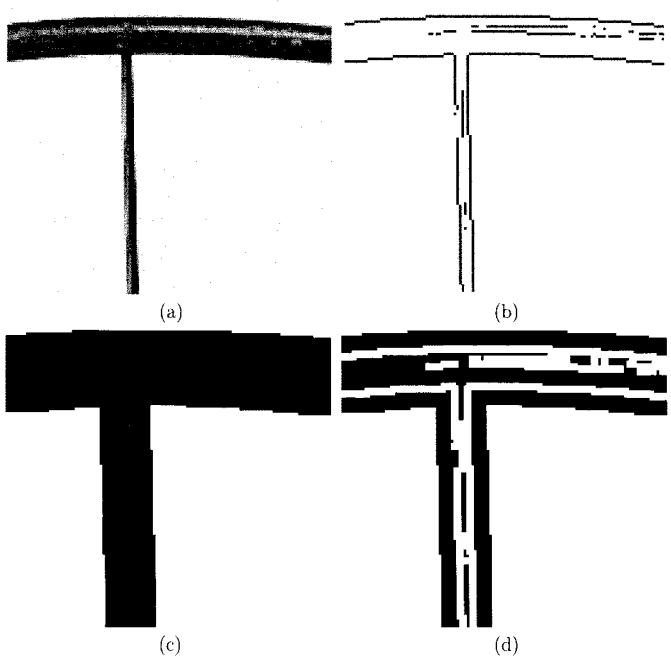

Fig. 8. Illustration of each step involved in VRM: (a) image; (b) edge map; (c) dilated edge map; and (d) VRM mask after the actual edges are excluded. VRM is the average local variance in the region indicated by the VRM mask.

TABLE III
VRM OF THE COMPRESSED IMAGES AND THE POST-PROCESSED IMAGES AT 0.125 BPP

Image	Compressed	Post-processed	
		k-means	*robust filter*
Aerial2	19.38	12.51	14.48
Bike	34.86	23.08	14.47
Café	51.01	46.08	30.14
Chart	18.13	13.87	14.36
Target	34.09	27.82	32.52
Tools	66.64	60.22	44.88
Woman	15.71	9.54	5.08

TABLE IV
VRM OF THE COMPRESSED IMAGES AND THE POST-PROCESSED IMAGES AT 0.0625 BPP

Image	Compressed	Post-processed	
		k-means	*robust filter*
Aerial2	45.50	42.26	29.07
Bike	37.02	27.51	16.56
Café	57.16	55.68	39.34
Chart	26.07	21.97	14.82
Target	79.07	75.08	78.18
Tools	55.97	50.66	37.20
Woman	25.95	20.54	11.09

TABLE V
DECODING TIME AND PROCESSING TIME OF THE ROBUST FILTER INCLUDING FILE I/O, ON A PC WITH DUAL 333 MHZ PENTIUM PROCESSORS AND 512 MBYTE MEMORY RUNNING WINDOWS NT 4.0

Image	Size	0.125 bpp		0.0625 bpp	
		Decode	Post-process	Decode	Post-process
Aerial2	[2048,2048]	4.74	3.31	4.59	4.16
Bike	[2560,2048]	5.83	5.88	5.59	6.52
Café	[2560,2048]	5.81	6.05	5.67	6.63
Chart	[2347,1688]	4.41	4.33	4.22	4.91
Target	[512,512]	0.48	0.28	0.45	0.31
Tools	[1200,1524]	2.14	1.79	2.05	2.09
Woman	[2560,2048]	5.80	5.08	5.58	5.75

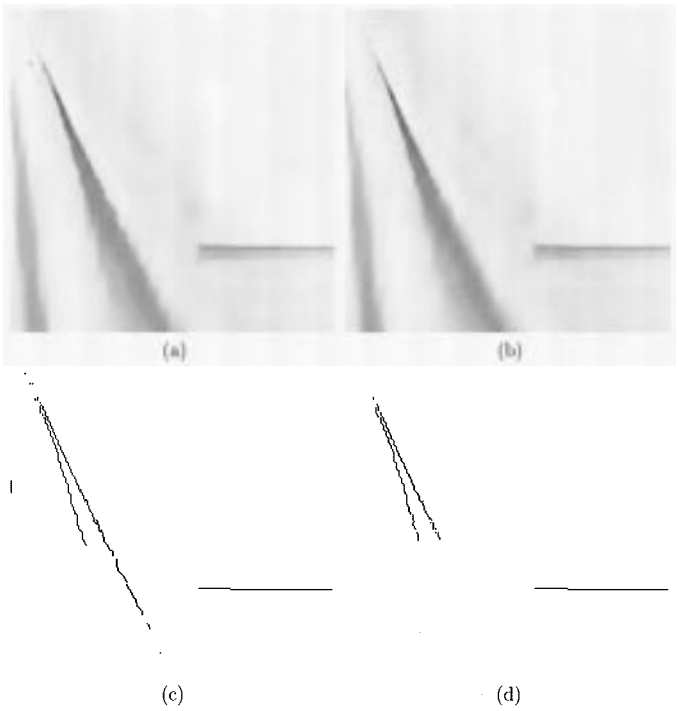

Fig. 9. Staircase effect, a part of the *Bike* image at 0.125 bpp (a) image post-processed by an existing method, (b) image post-processed by the proposed robust filter, (c) edge map of the image processed by an existing method, and (d) edge map of the image processed by the proposed robust filter.

store the intermediate results during the iteration. In the proposed method, the estimation depends only on the samples in a finite-size window. The algorithms can potentially be implemented with $w \times w$ byte buffers for the $[w \times w]$ size sliding window. Currently, the algorithms are implemented with $w \times N$ size buffers for an $[M \times N]$ image.

In some MAP estimation based approaches, prior knowledge of smoothness is modeled by the Gibbs distribution with a so-called *redescending function* [12] as a potential function. When a redescending function is used to model f, the feasible image set should be enforced strictly [3]. Otherwise, the artifact removal process might introduce another type of artifact called the *staircase effect* due to the property of redescending functions [20]–[22]. The staircase effect is a phenomenon in which a graded edge is broken into pieces during the restoration process. It is due to the fact that the redescending functions prefer step edges to graded edges. Consequently, a smoothly varying region is represented by a series of staircase-like step edges. Fig. 9 shows the comparison between one of the existing methods [23] and the robust filter. The algorithm in [23] approximates the projection onto the feasible image set with a nonlinear operation to reduce the computational complexity. The image in Fig. 9(a) shows the staircase effect of that algorithm. The edge map in Fig. 9(c) shows the false edges introduced by the staircase effect. In other parts of the image, the staircase effect is responsible for making post-processed images cartoon-like or painting-like. The image and the edge map in Fig. 9(b) and (d) show the result of the proposed robust filter. It shows better preservation of smooth regions by the proposed robust filter.

V. CONCLUSION

In this work, removal of the image ringing-artifact is posed as an ML parameter-estimation problem. The ML estimation viewpoint requires a new way to model the degraded image. We model the probability density of the degraded image as the one that peaks at the ringing-artifact-free flat surface. The ML estimate is obtained for the probability density estimated by the k-means algorithm with the number of flat surfaces determined by a criterion of merit. In contrast to the Bayesian viewpoint, the ML formulation leads to a simple approximation of the problem.

The proposed algorithm and its simplified version are applied to the ringing-artifact removal from JPEG2000 compressed images. The results show effective removal of the ringing artifact with significant reduction in computational effort in the three-cluster case. The simplified algorithm has the advantage over many post-processing algorithms because of its good performance and simplicity.

REFERENCES

[1] R. L. de Queiroz, T. Q. Nguyen, and K. R. Rao, "GenLOT: generalized linear-phase lapped orthogonal transform," *IEEE Trans. Signal Processing*, vol. 44, pp. 497–507, Mar. 1996.
[2] T. P. O'Rourke and R. L. Stevenson, "Improved image decompression for reduced transform coding artifacts," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp. 490–499, Dec. 1995.
[3] S. Yang, "Still image and video sequence coding artifact removal," Ph.D. dissertation, Univ. of Wisconsin, Madison, 2000.
[4] ——, "Blocking effect removal using robust statistics and line process," in *Proc. IEEE Int. Workshop on Multimedia Signal Processing*, 1999, pp. 315–320.
[5] J. A. Hartigan, "A k-means clustering algorithm," *Appl. Statist.*, vol. 28, pp. 100–108, 1979.

[6] D. L. Davies and D. W. Bouldin, "A cluster seperation measure," *IEEE Trans. Pattern Recognit. Machine Intell.*, vol. 1, no. 2, pp. 224–227, Apr. 1979.
[7] *Coding of Still Pictures*, ISO/ IEC/ JTC1/ SC29/ WG1, 1999.
[8] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, "Projection-based spatially adaptive reconstruction of block-transform compressed images," *IEEE Trans. Image Processing*, vol. 4, pp. 896–908, July 1995.
[9] J. K. Su and R. M. Mersereau, "Post-processing for artifact reduction in JPEG-compressed images," in *IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 4, 1995, pp. 2363–2366.
[10] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images," *IEEE Trans. Pattern Recognit. Machine Intell.*, vol. 6, pp. 721–741, Nov. 1984.
[11] P. J. Huber, *Robust Statistical Procedures*, 2nd ed. Philadelphia, PA: Siam, 1977.
[12] D. C. Hoaglin, F. Mosteller, and J. W. Tukey, *Understanding Robust and Exploratory Data Analysis*. New York: Wiley, 1983.
[13] J. Zhang, J. W. Modestino, and D. A. Langan, "Maximum-likelihood parameter estimation for unsupervised stochastic model-based image segmentation," *IEEE Trans. Image Processing*, vol. 3, pp. 404–419, July 1994.
[14] J. Zhang and J. W. Modestino, "A model fitting approach to cluster validation with application to stochastic model-based image segmentation," *IEEE Trans. Pattern Recognit. Machine Intell.*, vol. 12, pp. 1009–1017, Oct. 1990.
[15] T. N. Pappas and N. S. Jayant, "An adaptive clustering algorithm for image segmentation," in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, May 1988, pp. 1667–1670.
[16] P. A. Kelly, H. Derin, and K. D. Hartt, "Adaptive segmentation of speckled images using a hierarchical random field model," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1628–1641, Oct. 1988.
[17] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
[18] V. N. Vapnik, *The Nature of Statiscal Learning Theory*. New York: Springer-Verlag, 1995.
[19] S. H. Oguz, "Morphological post-filtering of ringing and lost data concealment in generalized lapped orthogonal transform based image and video coding," Ph.D. dissertation, Univ. of Wisconsin, Madison, 1999.
[20] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Recognit. Machine Intell.*, vol. 12, pp. 629–639, 1990.
[21] R. T. Whitaker and S. M. Pizer, "A multi-scale approach to nonuniform diffusion," *CVGIP: Image Understanding*, vol. 57, no. 1, pp. 99–110, Jan. 1993.
[22] M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger, "Robust anisotropic diffusion," *IEEE Trans. Image Processing*, vol. 7, pp. 421–432, Mar. 1998.
[23] M. Shen and C. C. J. Kuo, "Artifact removal in low bit rate wavelet coding with robust nonlinear filtering," in *IEEE Int. Workshop on Multimedia Signal Processing*, 1998, pp. 480–485.

Seungjoon Yang received the B.S. degree from the Seoul National University, Seoul, Korea, in 1990, and the M.S. and Ph.D. degrees from the University of Wisconsin at Madison in 1993 and 2000, respectively, all in electrical engineering.

He joined Digital Media System Lab. Corporate R&D Center, Samsung Electronics Co., Ltd., Suwon, Korea, in September 2000, as a Senior Researcher. His research interests are in digital and image processing, estimation theory, and multirate systems.

Yu-Hen Hu (S'80–M'83–SM'87–F'99) received the B.S.E.E. degree from National Taiwan University, Taipei, Taiwan, R.O.C., in 1976, and the M.S.E.E. and Ph.D. degrees in electrical engineering from University of Southern California at Los Angeles in 1980 and 1982, respectively.

From 1983 to 1987, he was an Assistant Professor in the Electrical Engineering Department, Southern Methodist University, Dallas, TX. He joined the Department of Electrical and Computer Engineering, University of Wisconsin at Madison in 1987 as an Assistant Professor (until 1989), and is currently an Associate Professor. His research interests include multimedia signal processing, artificial neural networks, fast algorithms and design methodology for application specific micro-architectures, and computer-aided design tools for VLSI using artificial intelligence. He has published more than 170 technical papers in these areas. His recent research interests focused on image and video processing and humancomputer interface.

Dr. Hu is a former Associate Editor (1988–1990) for the IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING in the areas of system identification and fast algorithms. He is currently Associate Editor of the *Journal of VLSI Signal Processing*. He is a founding member of the Neural Network Signal Processing Technical Committee of the IEEE Signal Processing Society and served as Chair from 1993 to 1996. He is a former member of the VLSI Signal Processing Technical Committee of the IEEE Signal Processing Society, and served as the society's Secretary during 1996–1998 .

Truong Q. Nguyen (S'85–M'89–SM'95) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, in 1985, 1986, and 1989, respectively.

He was with MIT Lincoln Laboratory, Cambridge, MA, from June 1989 to July 1994 as a member of Technical Staff. During the academic year 1993–1994, he was a Visiting Lecturer at MIT and an Adjunct Professor at Northeastern University, Boston, MA. During 1994–1998, he was with the Electrical and Computer Engineering Department, University of Wisconsin at Madison. He is now a Full Professor at Boston University, Boston, MA. He also co-organizes several workshops on wavelets at MIT. His research interests are in the theory of wavelets and filter banks and applications in image and video compression, telecommunications, bioinformatics, medical imaging and enhancement, and analog/digital conversion. He is the coauthor (with G. Strang) of the popular textbook *Wavelets and Filter Banks* (Cambridge, MA: Wellesley-Cambridge Press, 1997) and the author of several Matlab-based toolboxes on image compression, electrocardiogram compression, and filter-bank design. He also holds a patent on an efficient design method for wavelets and filter banks and several patents on wavelet applications, including compression and signal analysis.

Dr. Nguyen served as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1994 to 1996 and for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS from 1996–1997. He previously served on the Digital Signal Processing Technical Committee for the IEEE Circuits and Systems Society, and is currently the Series Editor (Digital Signal Processing) for Academic Press. He received the IEEE TRANSACTIONS ON SIGNAL PROCESSING Paper Award (Image and Multidimensional Processing) for the paper he co-wrote on linear-phase perfect-reconstruction filter banks in 1992, the National Science Foundation Career Award in 1995, and the Technology Award from Boston University for his invention on Integer DCT (with Ying Jui Chen and Soontorn Oraintara).

Damon L. Tull (S'96–M'97) received the B. S. degree in electrical engineering from the Rensselaer Polytechnic Institute, Troy, NY, and the M.S. and Ph.D. degrees in electrical engineering from Northwestern University, Evanston, IL, in 1994 and 1998, respectively.

From 1998 to 2001, he was an Assistant Professor with the University of Wisconsin–Madison. In 2000, he founded DVIP Multimedia Incorporated, Verona, WI, where he is the Chairman and CEO, responsible for guidging the development of new digital imaging technology. His research interests are in image sequence processing, object-based image recovery, color processing, and the restoration of images degraded by compression and acquisition distortions. He holds patents and has published in the areas of content interactive multimedia, stereoscopic vedio compression, motion estmation, image formation, and recovery.