Optimized Vector Quantization for Bayer Color Filter Array

M. Lakshmi, J. Senthil Kumar

Abstract—Digital cameras to reduce cost, use an image sensor to capture color images. Color Filter Array (CFA) in digital cameras permits only one of the three primary (red-green-blue) colors to be sensed in a pixel and interpolates the two missing components through a method named demosaicking. Captured data is interpolated into a full color image and compressed in applications. Color interpolation before compression leads to data redundancy. This paper proposes a new Vector Quantization (VQ) technique to construct a VQ codebook with Differential Evolution (DE) Algorithm. The new technique is compared to conventional Linde-Buzo-Gray (LBG) method.

Keywords—Color Filter Array (CFA), Biorthogonal Wavelet, Vector Quantization (VQ), Differential Evolution (DE).

I. INTRODUCTION

CFA is a distinctive hardware element in single-sensor imaging. It is placed atop a monochrome image sensor, usually a Charge-Coupled Device (CCD) or a Complementary Metal Oxide Semiconductor (CMOS) sensor, to get the image scene’s low-resolution color information. Every sensor cell has its spectrally selective filter and so acquired CFA data ensures a mosaic-like monochrome image [1]. A camera has to perform much processing when a digital image is recorded to provide a user with a viewable image. This includes the white balance adjustment, gamma correction, and compression.

A color image needs a minimum of 3 color samples at a pixel location. Computer images use red, green, and blue usually. A camera would require 3 separate sensors to make such measurements. Many cameras to reduce size and cost use one sensor array with a CFA [2] allowing only a spectrum’s one part spectrum to pass to the sensor resulting in only one color being measured at each pixel. In other words, the camera must estimate the 2 missing color values at each pixel. This is called demosaicking.

As an important step in digital cameras image processing, demosaicking has interested academia and industry. The simplest approach to demosaicking is treating color channels separately, filling the missing pixels in the two channels using a spatially invariant interpolation method called bilinear or bicubic interpolation [3]. CFA captures only one-third of color intensities with the full color image being generated through interpolation from captured data.

Nearest pixels are considered for bilinear interpolation. To get each pixel’s separate color channels, an image first passes through a Bayer filter which absorbs only one channel value per pixel. They are arranged in RGBG, GRGB, RGGB patterns, based on first pixel’s value, and later arranged as a lattice structure. Interpolation using color correlation ensures better performance as there is good correlation between R, G, and B components [4]. The output image suffers from alternating colors, which is often referred to as zippering, artifacts in Bayer pattern. They are unwanted features not appearing in the original image.

Bayer pattern is the commonly used CFA pattern, which measures the green image on a quincunx grid (half image resolution) and the red, blue images on rectangular grids (quarter image resolution). As peak sensitivity of Human Visual System (HVS) is in medium wavelengths, the green channel is measured at a greater sampling rate than the other two. This corresponds to the spectrum’s green portion [5].

More pertinent information is retained by compressing CFA data before interpolation, achieving higher compression rates and/or image quality. An application can be video conferencing where frame transmission rates and or image quality is increased. In addition, when considering image or video storage, the scheme allows more images storage or reconstruction of higher quality color images and videos [6].

There are two CFA image compression schemes categories: lossy and lossless. Lossy schemes compress CFA image by discarding visually redundant information and result in a higher compression ratio compared to lossless schemes. Different lossy compression techniques like vector quantization, transform followed by JPEG or JPEG 2000, discrete cosine transform, sub-band coding with symmetric short kernel filters, and low-pass filtering followed by JPEG 2000 or JPEG-LS (lossless mode) reduce data redundancy [7].

High-end photography applications like commercial poster production, original CFA images produce high quality full color images directly where CFA images lossless compression is necessary. Lossless image compression schemes like JPEG-LS and JPEG2000 can encode a CFA image with only fair results. Lossless compression is necessary where original CFA

M.Lakshmi is with D. D. C. S. M. Polytechnic College, Palacode, Dharmapuri, Tamilnadu, India (e-mail: lakshmimfa@gmail.com).
J. Senthil Kumar is with Sona College of Technology, Salem, Tamilnadu, India. e-mail: dr.senthilkumarj@rediffmail.com
images produce high quality full color image. A CFA image is first interpolated, after which demosaicking is performed to form full image before compression for storage in demosaicking schemes.

An alternative processing chain was proposed to move compression before demosaicking. As compression process occurs before demosaicking, digital cameras have simple designs and low power consumption, which leads to high demand for CFA images. Here, based on the color differences variances along different edges the missing green samples are first estimated. Missing red and blue components are projected based on interpolated green plane [8]. This algorithm effectively preserves the details in texture regions and simultaneously reduces color artifacts greatly.

Raw CFA images direct compression is inefficient as existing compression solutions are optimized for continuous tone images and is not so effective for mosaic-like images. Due to the image sensor’s non-uniform spectral sensitivity, different color channels pixels have different average intensity levels [9]. Therefore, intermixing pixels from various color channels leads to artificial discontinuity. To address this, advanced CFA compression schemes exploit various pre-processing operations before image encoding for compression tools optimal use.

Vector Quantization (VQ) is a scalar quantization simplification. Large dimension vectors produce transparency in quantization at a specific bit-rate chosen. In VQ, data is quantized as contiguous blocks called vectors instead of individual samples. VQ is a great tool following development of non-variational design algorithms like the Linde-Buzo-Gray (LBG) algorithm. Various images are divided into many K-dimension training vectors in LBG codebook generation. The representative codebook is generated from training vectors through clustering techniques [10]. However, besides spectral distortion, vector quantizer has its own limitations like computational complexity and memory requirements for codebooks searching and storing. For applications needing higher bit-rates computational complexity and memory requirements go up exponentially.

Codebook discovers transfistic relationships between people, bugs, code, test cases, specifications, and other related artifacts by mining any software repository [11]. It supports multiple information needs extensively with one data structure (directed graph) and an algorithm (regular language reachability). Code book optimization is NP-hard, and various meta-heuristic techniques were proposed in literature. Nearest-neighbor quantization, where a finite codebook is formed in feature space, after which a feature vector is encoded by index of nearest code vector, is a common way of discretizing continuous feature spaces [12]. Codebooks are formed in modern image and signal processing literature, not only for compressing high-dimensional data (traditional goal of quantization), but also to facilitate the next step of learning a statistical model for classification and inference.

VQ’s primary focus is determining a limited length of the proper codebook assigned to represent entire origin image data set as similar as possible. When an optimum codebook is learnt from training image, it preserves the highest possible fidelity, i.e. compressed image has no distortion with original image [13]. VQ’s encoding process with a selected code vector determines mapping of input data set to a finite codebooks collection. When image training patterns are marked with relative index to nearest code vector, encoding is over. This condition describes that a suitable codebook depicts entire image data. Normally, codebook size is much smaller than original image data set. So, the purpose of image compression is achieved. In decoding, associated sub-image is retrieved exactly by same codebook used in encoding phase. When every sub-image is fully reconstructed, decoding is over [14].

A most desirable VQ compression technology objective is to increase compression rate to ensure a higher fidelity. The higher the compression rate, i.e. (the lower the computational data), decreases the more the required memory and transmission channel bandwidth. VQ-based image compression system design is considered as an optimal codebook selection problem to satisfy the earlier 2 goals. Code vectors extraction is similar to classification of special groups of image patterns, which ensure good code vectors (cluster centers) as an approximation of the original image. Codebook generation is accomplished with a k-means like or Generalized Lloyd type clustering algorithms, but performance of gradient descent type learning methodology is sensitive to choice of initial codebook or parameter settings.

This investigation used a wavelet to decompose an image and coefficients are encoded using VQ whose code book is optimized using Differential Evolution algorithm. The rest of the paper is organized as: Section II reviews related work. Section III explains methodology, and Section IV discusses the results of experiments. Section V concludes the work.

II. RELATED WORKS

Bayer patterned color image compression based on wavelet transform, and all phase interpolation was realized by [15]. Results showed that the new algorithm suits Bayer image compression.

A modified H.264 intra prediction scheme to compress image and video data captured with a CFA was presented by [16]. The H.264 intra prediction modes were modified for green channel as green data was not sampled in a rectangular manner in Bayer pattern, a popular CFA design. The new method increased compression efficiency of I frames on green channel by 1.2 dB.

A new image compression method for multispectral imaging system presented by [17] encoded mosaicked multispectral image before interpolation. The decoder interpolated the image after decoding to reconstruct a full-resolution image. Simulation showed that the new method outperformed conventional methods, which encoded full-resolution image after interpolation, at all bit rates. A heuristic CFA design developed for the new method lead to lower bit rate.

An encoder for lossless compression of CFA data of a hierarchical predictor and context-adaptive arithmetic encoder were presented by [18]. The sub-sampled images were
encoded in order in hierarchical prediction; each sub-images contained one color component (red, green, or blue) in a Bayer CFA image. By gathering information from prediction processes like edge activity and neighboring errors, prediction error magnitude was estimated. Experimental results on simulated CFA images showed that the new method produced less bits per pixel than conventional lossless image compression methods and the recently developed lossless CFA compression algorithms.

A Compressed Sensing (CS) based solution to demosaicking images captured using general class of periodic CFAs was derived by [19]. It is known that periodic CFAs can effectively separate luminance and chrominance frequency bands. This ability was used to reduce artifacts associated with luminance-chrominance overlap at solver side. The modified compressive demosaicking method along with additional constraint that chrominance channels have smooth surfaces lead to improved results for periodic CFAs.

A lossy Bayer image compression pipeline achieved better than Set-Partitioning In Hierarchical Trees (SPIHT) coding as core of original data compression algorithm was presented by [20]. The method encoded/decoded images available in CFA format. After completing image coding, a demosaicking stage follows to produce a full color image. Simulation showed that pipelines which used SPIHT coding achieved great compression performance improvement and low artifacts compared to similar pipelines with a JPEG coder.

A lossy compression scheme for Bayer images was proposed by [21]. Wavelet coefficients shuffle exploring sub-band correlation suited it for zerotree coding. Improved SPIHT algorithm exploited data correlation in various directions under same resolution using one symbol to denote three zerotrees while SPIHT algorithm used 3 symbols. Simulation showed that the new compression scheme at low bit rate compared to JPEG2000 had better performance and low complexity. CPSNR was 2 db higher than JPEG2000 at bit rate 1 and was 0.7 db higher than JPEG2000 at bit rate 2.

The possibility of combining the characteristics of Bayer and FELICS to eliminate 3-color-channels transformation was researched by [22]. After studying context correlation in Bayer pattern by FELICS theory in an adjusted way, modifying original FELICS algorithm to suit Bayer pattern was addressed clearly. Usage of Bayer version, FELICS was validated at the end through experimental results.

A visual lossless compression of mosaic image based on wavelet sub-band substitute proposed by [23] consisted of three coding pipelines: green component lossless coder, red component near-lossless coder, and blue component near-lossless coder. So mosaic images were encoded at low complexity with a higher compression ratio. Compared to JPEG2000, the new method’s performance was better.

A H.264/AVC based compression-first approach that exploited color channels correlation was proposed by [24]. An inter-channel model described by a linear equation in RGB space was proposed. The R and B pixels were specifically predicted by a neighboring G pixel or demosaicked G pixel at same location. The new prediction scheme further reduced prediction error variance, and so the new method provided about 1.45dB gain on average and 2.2dB gain at maximum for green channel over the conventional compression-first scheme. For non-green channels, average gain was about 1.0dB and maximum was about 1.8dB.

A low-complexity integer-reversible spectral-spatial transform that allowed efficient lossless and lossy compression of CFA images was presented by [25]. The advantage of the new transform was its mapping pixel array values to a format that is directly compressed in a lossless, lossy, or progressive-to-loss less manner by existing typical image coders like JPEG 2000 or JPEG XR. So, no special codec design was needed to compress camera-raw data. Another advantage was that the new transform allowed mild compression of camera-raw data in a near lossless format, ensuring very high quality, offline post-processing, but where camera-raw files can be half the size of those of current camera-raw formats.

A new CFA image compression scheme to handle High Dynamic Range (HDR) representation introduced by [26] consisted of a series of preprocessing steps followed by a JPEG XR encoding module. An 8-directional edge-sensing mechanism operator reduced estimation error and preserved edge related information during color space conversion. The utilized YCoCg color space allowed a simplified pipeline implementation. The new method’s performance was validated through objective quality assessment metrics and complexity analysis. Extensive experimentation indicated that the new solution suited limited resource environments because of its low complexity and high performance.

III. METHODOLOGY

This section discusses the new differential evolution and wavelet based decomposition methods.

A. Biorthogonal Wavelet

Biorthogonal wavelet transforms are invertible transforms with perfect reconstruction properties and symmetric wavelet functions. Biorthogonal wavelets acquire these properties as they possess 2 sets of low-pass filters (for reconstruction), and high-pass filters (for decomposition). Each set is a dual of the other [27].

Biorthogonal wavelet transform is composed of the decomposition and reconstruction processes with 2 different wavelets $\psi$ and $\psi^\dagger$. $\psi$ is used in decomposition and $\psi^\dagger$ is used in reconstruction. $\psi$ and $\psi^\dagger$ are dual and orthogonal to each other. This relationship is termed biorthogonal. There are 2 scale functions $\phi$ and $\phi^\dagger$, which are also dual and orthogonal [28].

One is used in decomposition and the other in reconstruction. So, there are 4 filters in biorthogonal wavelet transform which are decomposition low-pass filter $\{h_0\}$, decomposition high-pass filter $\{g_0\}$, reconstruction low-pass filter $\{\tilde{h}_0\}$, and reconstruction high-pass filter $\{\tilde{g}_0\}$. Unlike orthogonal wavelet transform, reconstruction and
decomposition filters are different. The dual scaling and wavelet functions have these properties [29]:
1. They are zero outside a segment.
2. The calculation algorithms are maintained and are very simple.
3. Associated filters are symmetrical.
4. Functions used in calculations are easier to build numerically than those in Daubechies wavelets.

Scaling equations in scaling functions and wavelets reveal that a signal’s decomposition and reconstruction from one resolution to the next is implemented by perfect reconstruction filter banks [30].

\[ a[n] = a_0 \cdot h[2n] \]
\[ d[n] = a_0 \cdot g[2n] \]

With
\[ h[n] = h[-n] \] and \[ g[n] = g[-n] \]

B. Vector Quantization (VQ)

VQ is a method widely used for data compression because of its theoretical advantages over scalar quantization schemes. However, computational complexity of both codebook design and vector lookup during encoding is obviously a burden of its realization. As the codebook is pre-generated before encoding, vector look up efficiency is comparatively significant [31].

In VQ, the image to be encoded is segmented into a set of input image vectors. An important task for VQ is designing a good codebook needed as the reconstructed image depends on code words in this codebook. The generated codebook is stored in a text file for vhdl file handling or data array in vhdl code.

C. Proposed Differential Evolution (DE) for VQ

Differential Evolution (DE) is a population-based search strategy similar to standard evolutionary algorithms. The difference is in reproduction step where offspring are created from 3 parents using an arithmetic cross-over operator. DE is defined for individual’s floating-point representation. Differential evolution does not use a mutation operator based on a probability distribution function but introduces a new arithmetic operator, which depends on differences between randomly selected pairs of individuals.

DE, compared to other evolutionary algorithms, reserves population-based global search strategy using a simple mutation operation of differential equation and one-on-one competition. It thus reduces the operation’s genetic complexity. Simultaneously, DE’s specific memory capacity enables it to dynamically track current search to adjust search strategy with a strong global convergence and robustness. So it suits solving complex environments in optimization problems [32].

(1) Initialize the number of population NP,
the maximum number of evolution Max inter, the scale factor and cross-factor.
(2) Initialize the population pop.
(3) Follow the DE policy enforcement options, and produce a new generation of individual:
(a) mutation operation;
(b) crossover operation;
(c) selection operation.
(4) Until the termination criterion is met.

DE evolves a population of NP n-dimensional individual vectors, i.e. solution candidates, \( X_i = (x_{i1},...,x_{in}) \) \( \in \mathbb{S} \), \( i = 1,...,NP \), from one generation to next. An initial population should cover entire parameter space by randomly distributing an individual’s parameter with uniform distribution between prescribed upper/lower parameter bounds \( x_{ij}^u \) and \( x_{ij}^l \). At each generation \( G \), DE uses mutation and crossover operations to produce a trial vector \( U_{i,G} \) for every individual.
vector \( X_{Gv} \), also called target vector, in current population [33].

For optimizing of VQ, the codebook must find optimal N-partition of the M training vectors with minimum distortion. The DE is used to find the optimal partition cut. The initial population is initialized randomly and the fitness of the individual is evaluated. The fitness is based on the measure of the performance of the codebook. The mutation, crossover and selection operators further improve the individuals in each iteration. This process is executed for a fixed number of generations and the best solution is chosen.

IV. RESULTS AND DISCUSSION

Experiments were also conducted using LBG and the proposed DE with VQ before and after denoising. The average BPP, MSE, CR and PSNR of the images with proposed method are shown in Tables I-IV.

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>EXPERIMENTAL RESULTS OF BPP WITH AND WITHOUT DENOISING</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>without denoising</td>
</tr>
<tr>
<td>LBG</td>
<td>5.146</td>
</tr>
<tr>
<td>DE</td>
<td>5.083</td>
</tr>
</tbody>
</table>

From Fig. 3, it can be observed that the DE method reduces BPP by 1.23\% without denoising and by 0.56\% with denoising when compared with LBG method.

<table>
<thead>
<tr>
<th>TABLE II</th>
<th>EXPERIMENTAL RESULTS OF MSE WITH AND WITHOUT DENOISING</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>without denoising</td>
</tr>
<tr>
<td>LBG</td>
<td>6.5203333</td>
</tr>
<tr>
<td>DE</td>
<td>7.6368333</td>
</tr>
</tbody>
</table>

From Fig. 4, it is observed that when denoising is applied, the average MSE is decreased by 14.21\% for proposed DE and by 15.78\% when compared to LBG without denoising.

<table>
<thead>
<tr>
<th>TABLE III</th>
<th>EXPERIMENTAL RESULTS OF CR WITH AND WITHOUT DENOISING</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>without denoising</td>
</tr>
<tr>
<td>LBG</td>
<td>64.3266667</td>
</tr>
<tr>
<td>DE</td>
<td>63.54</td>
</tr>
</tbody>
</table>

From Fig. 5, it can be observed when denoising is applied, the DE method reduces average CR by 0.56\% and by 1.23\% without denoising when compared with LBG method.

<table>
<thead>
<tr>
<th>TABLE IV</th>
<th>EXPERIMENTAL RESULTS OF PSNR WITH AND WITHOUT DENOISING</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>without denoising</td>
</tr>
<tr>
<td>LBG</td>
<td>39.9833333</td>
</tr>
<tr>
<td>DE</td>
<td>39.3366667</td>
</tr>
</tbody>
</table>

From Fig. 6, it is observed that the proposed DE increased average PSNR by 1.64\% without denoising and by 1.47\% with denoising when compared to LBG.
Fig. 6 Average PSNR

Most digital cameras to reduce cost use one image sensor to capture color images. A Bayer CFA is coated over a sensor in such cameras to record only one of the three color components at every pixel location. A CFA image is first interpolated through a demosaicking process to form a full color image before being compressed for storage. This paper proposed a new Vector Quantization (VQ) technique, which finds better solution in VQ space using the Differential Evolution (DE) algorithm. Results showed that the DE method reduces BPP by 0.56% with denoising compared to the LBG method. DE also increased average PSNR by 1.47% with denoising compared to LBG.

V. CONCLUSION

REFERENCES


