Collaborative Adaptive Down-Sampling And Upconversion - An Approach For Image Compression

Tahir Naquash H. B.
M.Tech Student, Dept of CS&E
MCE, Hassan.

Dr. V. N. Narayana
Professor, Dept Of CS &E
MCE, Hassan.

ABSTRACT:

In this paper, we are going to use a practical approach of uniform down sampling in image space and yet making the sampling adaptive by spatially varying, directional low-pass pre-filtering. The resulting down-sampled pre-filtered image remains a conventional square sample grid, and, thus, it can be compressed and transmitted without any change to current image coding standards and systems. The decoder first decompresses the low-resolution image and then up-converts it to the original resolution in a constrained least squares restoration process, using a 2-D piecewise autoregressive model and the knowledge of directional low-pass pre-filtering. The proposed compression approach of collaborative adaptive down-sampling and upconversion (CADU) outperforms JPEG 2000 in PSNR measure at low to medium bit rates and achieves superior visual quality, as well. The superior low bit-rate performance of the CADU approach seems to suggest that over-sampling not only wastes hardware resources and energy, and it could be counterproductive to image quality given a tight bit budget.

1. INTRODUCTION

The term digital image refers to processing of a two dimensional picture by a digital computer. In a broader context, it implies digital processing of any two dimensional data. A digital image is an array of real or complex numbers represented by a finite number of bits. An image given in the form of a transparency, slide, photograph or an X-ray is first digitized and stored as a matrix of binary digits in computer memory. This digitized image can then be processed and/or displayed on a high-resolution television monitor. For display, the image is stored in a rapid-access buffer memory, which refreshes the monitor at a rate of 25 frames per second to produce a visually continuous display [1].

1.1 THE IMAGE PROCESSING SYSTEM

A typical digital image processing system is given in fig.1.1

1.1.1 DIGITIZER

A digitizer converts an image into a numerical representation suitable for input into a digital computer. Some common digitizers are
1. Microdensitometer
2. Flying spot scanner
3. Image dissector
4. Videocon camera
5. Photosensitive solid-state arrays.
1.1.2 IMAGE PROCESSOR
An image processor does the functions of image acquisition, storage, preprocessing, segmentation, representation, recognition and interpretation and finally displays or records the resulting image. The following block diagram gives the fundamental sequence involved in an image processing system.

![Block Diagram of Fundamental Sequence involved in an image Processing system](image.png)

As detailed in the diagram, the first step in the process is image acquisition by an imaging sensor in conjunction with a digitizer to digitize the image. The next step is the preprocessing step where the image is improved being fed as an input to the other processes. Preprocessing typically deals with enhancing, removing noise, isolating regions, etc. Segmentation partitions an image into its constituent parts or objects. The output of segmentation is usually raw pixel data, which consists of either the boundary of the region or the pixels in the region themselves.

Representation is the process of transforming the raw pixel data into a form useful for subsequent processing by the computer. Description deals with extracting features that are basic in differentiating one class of objects from another. Recognition assigns a label to an object based on the information provided by its descriptors. Interpretation involves assigning meaning to an ensemble of recognized objects. The knowledge about a problem domain is incorporated into the knowledge base. The knowledge base guides the operation of each processing module and also controls the interaction between the modules. Not all modules need be necessarily present for a specific function. The composition of the image processing system depends on its application. The frame rate of the image processor is normally around 25 frames per second.

1.1.3 DIGITAL COMPUTER
Mathematical processing of the digitized image such as convolution, averaging, addition, subtraction, etc. are done by the computer.

1.1.4 MASS STORAGE
The secondary storage devices normally used are floppy disks, CD ROMs etc.

1.1.5 HARD COPY DEVICE
The hard copy device is used to produce a permanent copy of the image and for the storage of the software involved.

1.1.6 OPERATOR CONSOLE
The operator console consists of equipment and arrangements for verification of intermediate results and for alterations in the software as and when required. The operator is also capable of checking for any resulting errors and for the entry of requisite data.

2. RELATED WORK
This paper surveys an emerging theory which goes by the name of “compressive sampling” or “compressed sensing,” and which says that this conventional wisdom is inaccurate. Perhaps surprisingly, it is possible to reconstruct images or signals of scientific interest accurately and sometimes even exactly from a number of samples which is far smaller than the desired resolution of the image/signal, e.g. the number of pixels in the image. It is believed that compressive sampling has far reaching implications. For example, it suggests the possibility of new data acquisition protocols that translate analog information into digital form with fewer sensors than what was considered necessary. This new sampling theory may come to underlie procedures for sampling and...
compressing data simultaneously. In this short survey, we provide some of the key mathematical insights underlying this new theory, and explain some of the interactions between compressive sampling and other fields such as statistics, information theory, coding theory, and theoretical computer science [1].

The performance of image interpolation depends on an image model that can adapt to non-stationary statistics of natural images when estimating the missing pixels. However, the construction of such an adaptive model needs the knowledge of every pixels that are absent. This paper resolves this dilemma by a new piecewise 2D autoregressive technique that builds the model and estimates the missing pixels jointly. This task is formulated as a non-linear optimization problem. Although computationally demanding, the new non-linear approach produces superior results than current methods in both PSNR and subjective visual quality. Moreover, in quest for a practical solution, it breaks the non-linear optimization problem into two sub problems of linear least-squares estimation. This linear approach proves very effective in our experiments [2].

JPEG 2000, the new ISO/ITU-T standard for still image coding, has recently reached the International Standard (IS) status. Other new standards have been recently introduced, namely JPEG-LS and MPEG-4 VTC. This paper provides a comparison of JPEG 2000 with JPEG LS and MPEG-4 VTC, in addition to older but widely used solutions, such as JPEG and PNG, and well established algorithms, such as SPIHT. Lossless compression efficiency, fixed and progressive lossy rate-distortion performance, as well as complexity and robustness to transmission errors, are evaluated. Region of Interest coding is also discussed and its behavior evaluated. Finally, the set of provided functionalities of each standard is also evaluated. In addition, the principles behind each algorithm are briefly described. The results show that the choice of the “best” standard depends strongly on the application at hand, but that JPEG 2000 supports the widest set of features among the evaluated standards, while providing superior rate-distortion performance in most cases [3].

3. PROBLEM STATEMENT
We propose a new, standard-compliant approach of coding uniformly down-sampled images, which outperforms JPEG 2000 in both PSNR and visual quality at low to modest bit rates.

4. UNIFORM DOWN-SAMPLING WITH ADAPTIVE DIRECTIONAL PREFILTERING
Out of practical considerations, we make a more compact representation of an image by decimating every other row and every other column of the image. This simple approach has an operational advantage that the down-sampled image remains a uniform rectilinear grid of pixels and can readily be compressed by any of existing international image coding standards. To prevent the down-sampling process from causing aliasing artifacts, it seems necessary to low-pass prefiler an input image to half of its maximum frequency fmax. However, on a second reflection one can do somewhat better. In areas of edges, the 2-D spectrum of the local image signal is not isotropic. Thus, we seek to perform adaptive sampling, within the uniform down-sampling framework, by judiciously smoothing the image with directional low-pass prefiltering prior to down-sampling.

In addition, the directional low-pass filter design serves two other purposes: 1) most efficient packing of signal energy in presence of edges; 2) preservation of subjective image quality for the edge is an important semantic construct. Moreover, as we will see in the next section, the use of low-pass prefiltering establishes sample relations that play a central role in the decoding process of constrained least squares up conversion.

Many implementations of directional lowpass preilters are possible. For instance, the following directional low-pass prefilter can be used:

\[ h_\theta(i,j) = m \text{sinc} \left( \frac{i \cos \theta + j \sin \theta}{s_i} \right) \]

\[ \times \text{sinc} \left( \frac{-i \sin \theta + j \cos \theta}{s_j} \right) \psi(i,j) \]
where \( m \) is the normalization factor to keep the filter in unit energy, and \( \Psi(i,j) \) is a window function (such as the window function). The parameters \( S_i \) and \( S_j \) are

\[
S_i = \frac{w_H(\theta)}{2\pi}, \quad S_j = \frac{w_L(\theta)}{2\pi}.
\]

Despite its simplicity, the CADU compression approach via uniform down-sampling is not inherently inferior to other image compression techniques in rate-distortion performance, as long as the target bit rate is below a threshold.

5. CONSTRAINED LEAST SQUARES UPCONVERSION WITH AUTOREGRESSIVE MODELING

In this section, we develop the decoder of the CADU image compression system. Let \( I_\downarrow \) be the decompressed \( M/2 \times N/2 \) subsampled image, and \( I \) be the original \( M \times N \) image. The sample relation between \( I \) and \( I_\downarrow \) is illustrated by Fig. 3. The decoder upconverts \( I_\downarrow \) to the original resolution of \( I \) by a constrained least squares reconstruction algorithm. The upconversion is based on a piecewise autoregressive image model and on the deconvolution of the directional low-pass prefiltering.

First, we introduce a suitable image model to aid the image recovery process. Our joint design of encoder and decoder gives a priority to the reconstruction of significant edges. There are two reasons for this. One is that human visual system is sensitive to phase errors in edge reconstruction. The other is that an edge can be down sampled along its direction and still reconstructed via directional interpolation. A good model for edges of sufficient scale is one of piecewise autoregressive (PAR) process [2], [3].

Fig. 3. Relationship between the down-sampled

\[
I(\hat{i},\hat{j}) = \sum_{(m,n) \in W_{i,j}} \alpha_{m,n} I(\hat{i} + m, \hat{j} + n) + v_{i,j}
\]

where \( W_{i,j} \) is a local window centered at pixel \((i,j)\), and \( v_{i,j} \) is a random perturbation independent of pixel location \((i,j)\) and the image signal. The term \( v \) accounts for both the fine-scale randomness of image signal and measurement noises.

The parameters \( \alpha_{m,n} \) of the autoregressive model specify the direction and amplitude of edges. If the dominant feature in the window \( W_{i,j} \) is an edge, then the parameters \( \alpha_{m,n} \) do not change in the window, and the PAR model fits both down-sampled image \( I_\downarrow \) and the original image \( I \) in \( W_{i,j} \). Therefore, the upconversion process can learn the edge structure by fitting samples of \( I_\downarrow \) in \( W_{i,j} \) to the parametric model. The validity of the proposed PAR image model hinges on the assumption of piecewise stationarity of the image signal. The assumption is acceptable for edges of sufficiently large scale, although a natural image generally has nonstationary statistics across different image segments.

To simplify notations, from now on we use a single index to identify 2-D pixel locations, and denote the pixels in \( I_\downarrow \) and \( I \) by \( y_i \) and \( x_i \) respectively. The 8-connected neighbors and 4-connected neighbors of an original pixel \( x_i \) in \( I \) are labeled by \( x^x_i \) and \( x^+_{i,t} \) \( t = 0,1,2,3 \). Similarly, the 8-connected and 4-connected neighbors of down-sampled pixel \( y_i \) in \( I_\downarrow \) are denoted as \( y^x_i \) and \( y^+_{i,t} \) \( t = 0,1,2,3 \).

Now we are ready to state the task of upconverting \( I_\downarrow \) to \( I \) as the following constrained least squares problem:
\[
\min_\mathbf{x} \left\{ \xi^- \sum_{i \in \mathbf{W}} \left[ x_i - \sum_{0 \leq n \leq 3} a_n x_{i-n} \right]^2 + \xi^+ \sum_{i \in \mathbf{W}} \left[ x_i - \sum_{0 \leq n \leq 3} b_n x_{i-n} \right]^2 \right\}
\]
subject to \[\| \mathbf{x} + \mathbf{h} - \mathbf{y} \|^2 = \| \mathbf{m}_W (r) \|^2 \] for \( \mathbf{x} \in \mathbf{W} \).

Where \( \mathbf{a} = (a_0, a_1, a_2, a_3) \) and \( \mathbf{b} = (b_0, b_1, b_2, b_3) \) are two sets of autoregressive coefficients, \( \xi^- \) and \( \xi^+ \) are the two corresponding least squares weights to be clarified shortly. The constraint \( \| \mathbf{x} + \mathbf{h} - \mathbf{y} \| = \| \mathbf{m}_W (r) \| \) is in accordance with the prefiltering and compression operations at the encoder side, and by which the decoder collaborates with the encoder.

We formulated the constrained least squares problem using two PAR models of order 4 each: the model of parameters \( \mathbf{a} \) and the model of parameters \( \mathbf{b} \). The two PAR models characterize the axial and diagonal correlations, respectively, as depicted in Fig. 4. These two models act, in a predictive coding perspective, as noncausal adaptive predictors. This gives rise to an interesting interpretation of the CADU decoder: adaptive noncausal predictive decoding constrained by the prefiltering operation of the encoder.

### IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

Extensive experiments were carried out to evaluate the proposed image coding method, in both PSNR and subjective quality. We compared the CADU method with the adaptive downsampling-based image codec proposed by Lin and Dong [7]. The latter was reportedly the best among all previously published downsampling-interpolation image codecs [5], [10] in both objective and subjective quality. Note that all existing image codecs of this type were developed for DCT-based image compression, whereas the CADU method is applicable to wavelet-based codecs as well. Therefore, we also include in our comparative study JPEG 2000, the quincunx coding method [9], and the method of uniform down-sampling at the encoder and bicubic interpolation at the decoder. The bicubic method in the comparison group and the CADU method used the same simple encoder: JPEG 2000 coding of uniformly down-sampled prefiltered image. The difference is in the upconversion process: the former method performed bicubic image interpolation followed by a deconvolution step using Weiner filter to reverse the prefiltering,
instead of solving a constrained least squares image restoration problem driven by autoregressive models as described in the proceeding section.

Although the proposed CADU method favors the reconstruction of edges, we chose, for fairness and generality of our comparative study, a large set of test images of various scene positions. Here, we report experimental results for four representative images, which represent all common image features in balance: edges of all scales, smooth regions, and granular textures.

Table II lists the PSNR results of seven methods: DCT-based old JPEG standard (column JPEG), the method of Lin and Dong [7] (the second column), the CADU method coupled with DCT-based JPEG (column CADU-JPG) JPEG 2000 (column J2K), JPEG 2000 coupled with uniform downsampling and bicubic interpolation (column Bicubic-J2K), the quincunx method [9] (column Quincunx), and the CADU method coupled with JPEG 2000 (column CADU-J2K). The results are tabulated against various bit rates from 0.1 bpp to 0.3 bpp. For the first two methods of the group, some table entries at very low bit rates are “N/A” because the DCT-based JPEG cannot even operate at such low rates for
the image tested.

The PSNR comparison between Lin-Dong’s method and the CADU-JPG method is somewhat mixed. CADU-JPG has a small advantage over Lin-Dong’s method in most cases, but the former loses to the latter by small margin for test images Flower and Bike when the rate is equal to and greater than 0.25 bpp. Although these two methods outperform old JPEG without down-sampling, they both produced significantly lower PSNR than wavelet-based JPEG 2000 without down-sampling.

Obviously, one should use JPEG 2000 when the bit budget is low in practice. For the state-of-the-art in low bit rate image compression, the reader should pay closer attention to the results of the wavelet group in Table II. At low rates, the CADU-J2K method achieves up to 0.5 dB higher PSNR than JPEG 2000. This improvement is appreciable given that JPEG 2000 is highly regarded for its outstanding performance at low rates [11]. Among the four competing methods in the wavelet group, the bicubic interpolation method has the lowest PSNR in most cases. Given that the CADU-J2K and bicubic interpolation methods use the same prefilters and the same JPEG 2000 encoder, the performance gap between the two manifests the efficacy of least squares noncausal predictive decoding constrained by adaptive directional low-pass prefiltering.

The quincunx coding method also outperforms JPEG 2000 at low to modest bit rates, but it requires a much more expensive, nonstandard encoder.

Next, let us assess the subjective quality of the methods evaluated. Fig. 5 presents the decoded images by different methods at bit rate 0.2 bpp. First, we notice that the wavelet-based methods have superior visual quality to the DCT-based methods, which is consistent with the PSNR comparison results in Table II. In the wavelet group, the CADU-J2K method produces the visually most pleasing images. At low bit rates, both JPEG 2000 and the bicubic interpolation method produce objectionable visual artifacts (e.g., jaggies and ringings) in edge areas, whereas the CADU-J2K method is largely free of those defects. Even when the bit rate gets higher and JPEG 2000 starts to have higher PSNR than the CADU-J2K method, its visual quality still appears inferior, as demonstrated by examples in Fig. 6. The superior visual quality of the CADU-J2K method is due to the good fit of the piecewise autoregressive model to edge structures and the fact that human visual system is highly sensitive to phase errors in reconstructed edges. We believe that the CADU-J2K image coding approach of down-sampling with directional pre-filtering at the encoder and edge-preserving upconversion at the decoder offers an effective and practical solution for subjective image coding.

Some viewers may find that JPEG 2000 produces somewhat sharper edges compared with CADU-J2K, although at the expense of introducing more and worse artifacts. However, one can easily tip the quality balance in visual characteristics to favor CADU-J2K by performing an edge enhancement of the results of CADU-J2K. Fig. 7 presents some sample results of JPEG 2000 and CADU-J2K at the bit rate of 0.2 bpp after edge enhancement. For better judgement these images should be compared with their counterparts in Fig. 5. As expected, the high-pass operation of edge enhancement magnifies the structured noises accompanying edges in images of JPEG 2000. In contrast, edge enhancement sharpens the images of CADU-J2K without introducing objectionable artifacts, which further improves the visual quality.

The CADU-J2K decoder has much higher complexity than the decoder based on bicubic interpolation. A close inspection of the reconstructed images by the CADU-J2K decoder and the bicubic method reveals that the two methods visually differ only in areas of edges. Therefore, an effective way of expediting the CADU-J2K decoder is to invoke least squares noncausal predictive decoding, which is the computation bottleneck of CADU, only in regions of high activity, and resort to fast bicubic interpolation in smooth regions. If a decoder is
severely constrained by computation resources, it can perform bicubic interpolation everywhere in lieu of the CADU restoration process.

Fig 6: Waiting for input for decomposition level

Fig 7: Decompressed image and Compression Ratio
V. CONCLUSION

We proposed a new, standard-compliant approach of coding uniformly down-sampled images, which outperforms JPEG 2000 in both PSNR and visual quality at low to modest bit rates. This success is due to the novel upconversion process of least square noncausal predictive decoding, constrained by adaptive directional low-pass prefiltering. Our findings suggest that a lower sampling rate can actually produce higher quality images at certain bit rates. By feeding the standard methods downsampled images, the new approach reduces the workload and energy consumption of the encoders, which is important for wireless visual communication.

REFERENCES


About the Authors:


2. Dr. V. N. Narayana, Professor, Dept of CS&E He is serving MCE since 1983. His areas of interest are A.I, NLP, Machine translation, Compilers and Operating Systems.