The Precedence of the Work of Said and Pearlman over Ricoh's CREW

Amir Said and William A. Pearlman

Introduction

The objective of this document is to establish without doubt the similarities between the papers "Reversible image compression via multiresolution representation and predictive coding," [1] published in November of 1993, and "CREW: Compression with reversible embedded wavelets," [2] published in March of 1995.

As authors of the first paper we want make clear that, as the first to publish, we deserve the credit for the discoveries contained on it. Unfortunately, this is necessary because the authors of the CREW paper certainly (as we show below) knew about our work, but failed to yield the smallest recognition. We are aware that simultaneous scientific discoveries occur all the time. However, it is also well known to us that the euphoria of each new discovery is often followed by the anxiety of thinking: "Haven't anyone published something similar or (God forbid) equal?" The reason for anxiety is that, independently of the brilliance of our derivation, those that published it first deserve the main credit. This purposeful concealment of information annoyed us because it can only indicate: (i) extreme arrogance; (ii) fear of a fair comparison; (iii) existence of hidden interests; (iv) all above.

We deeply regret to have to publicly discuss this subject. However, we became tired of being frequently asked by researchers: "Isn't the work you're doing VERY similar to CREW?" Absurdly, this sounds like WE are copying their work! So, what is at stake here is not simply the vanity of being called the first, but the necessity of making the truth clear and well known.

The Compression Method

In [1] we propose a compression scheme that is very efficient for lossy and for lossless compression, and furthermore can implement both forms of compression simultaneously in a seamless fashion. In our method the same bit stream contains, sequentially, all the information for good lossy reproductions at low bit rates, and for gradually and efficiently reducing the distortion up to perfect reconstruction. Presently there are comparable methods for lossy and for lossless compression, and also some that can do both, but not simultaneously.

To achieve what we believe was a significant advance in state-of-the-art at that time, we proposed two innovations:

(a) An integer-valued multiresolution representation, similar to the subband representation, but that deals carefully with the sequence of arithmetic operations and truncations to make it efficient for lossless compression.

(b) The use of the (then recently proposed) Embedded Zerotree Wavelet (EZW) method adapted to use the transform of (a), normalized to be nearly unitary.

Of course, the full implementation requires a description of several aspects, but the distinguishing properties of the new compression method all come from these two new proposals and their integration in a single compression method.

The Image Transformation

We next show that the transformation used by CREW [2] is not similar, but indeed *equal* to a special case of our S+P transform [1] (which was proposed in a much more general context). Both transformations are described as an extension of the S Transform. In our paper the sequence $x[n]$ represents a one-dimensional sequence of image pixels, and equation (1) of [1] describes the truncated lowpass components as used in the S Transform

$$l[n] = \lfloor (c[2n] + c[2n+1])/2 \rfloor \,. \tag{1}$$

In the same way the CREW paper [2] uses $x(n) \equiv c[n]$, and defines in equations (3) and (6).

$$r(n) = \left\lfloor \frac{x(2n) + x(2n+1)}{2} \right\rfloor . \tag{2}$$

Referring to the TS (two-six) subband transformation, with lowpass filter taps $(1,1)/\sqrt(2)$ and $(-1,-1,8,-8,1,1)/8\sqrt{2}$, the CREW paper then defines the highpass components in eq. (6) by

$$d(n) = \left\lfloor \frac{-\left\lfloor \frac{x(2n)+x(2n+1)}{2} \right\rfloor + 4(x(2n+2) - x(2n+3)) + \left\lfloor \frac{x(2n+4)+x(2n+5)}{2} \right\rfloor}{4} \right\rfloor \tag{3}$$

Note that just using integer tap values on the filters does not make a subband transformation good for lossless compression. What is needed is the use of truncations to eliminate redundancy in the least-significant bits. However, in the CREW paper there is absolutely no indication of how they identified the right sequence of truncations.

In our paper we can "bactrack" the same equation from Table 1 (predictor "A": $\alpha_1 = \alpha_2 = 1/4, \alpha_3 = 0$), eqs. (5), (4), (3), and (1):

$$h_d[n] \;\; = \;\; h[n] - \left\lfloor \hat{h}[n] \right\rfloor$$

$$
\begin{aligned}
&= h[n] - \lfloor \alpha_1 \, \Delta l[n] + \alpha_2 \, \Delta l[n+1] - \alpha_3 h[n+1] \rfloor \\
&= c[2n] - c[2n+1] - \left\lfloor \frac{\Delta l[n] + \Delta l[n+1]}{4} \right\rfloor \\
&= c[2n] - c[2n+1] - \left\lfloor \frac{(l[n-1] - l[n]) + (l[n] - l[n+1])}{4} \right\rfloor \\
&= c[2n] - c[2n+1] - \left\lfloor \frac{l[n-1] - l[n+1]}{4} \right\rfloor \\
&= c[2n] - c[2n+1] - \left\lfloor \frac{\left\lfloor \frac{c[2n-2]+c[2n-1]}{2} \right\rfloor - \left\lfloor \frac{c[2n+2]+c[2n+3]}{2} \right\rfloor}{4} \right\rfloor \\
&= \left\lfloor \frac{-\left\lfloor \frac{c[2n-2]+c[2n-1]}{2} \right\rfloor + 4(c[2n] - c[2n+1]) + \left\lfloor \frac{c[2n+2]+c[2n+3]}{2} \right\rfloor}{4} \right\rfloor
\end{aligned}
$$

So, $l[n] \equiv r(n)$ and $h_d[n] \equiv d(n-1)$, showing that the transformations are identical.

We used a more complex notation because: (i) we explain exactly when the truncations can and should be used; (ii) we have a much more general formulation, and analyzed and tested the choice of different parameters.

The possibility of using fast bit shifts instead of multiplications/division is mentioned in both [1] and [2].

Entropy-Coding the Transformed Image

The special entropy-coding method for the transformed image—an adapted version of the EZW [3] method—is proposed by... both papers. The papers present few details about their respective implementations, but [1] cites paper [4] where the algorithm is described.

In fact, the paper [4] was the first to show results superior to Shapiro's, and also the first to explain that the EZW method is similar to a bit-plane transmission. Thus, it follows easily that it is suited for lossless compression. We wonder if J. Shapiro discovered this potential for his method, but anyway at that time he did not have a good transformation for lossless compression,

Thus, we were first to effectively test the EZW for lossless compression (which was then known only for lossy compression), and (equally important) propose the efficient transformation that was required for that type of application. A very important implementation aspect—the necessity of using an implicit form of scaling to have a nearly unitary transformation—is already mentioned in [1]. We had shown that the loss due to having embedded coding is insignificant or very small.

One further implementation detail was missing in our paper: the observation that EZW may become inefficient when coding the least-significant bits. For that reason in CREW it is suggested to use a second entropy-coding method (which they call Horizon) to code the least-significant bits. Our coding algorithm [1] is less sensitive to this problem, but we used the same approach. Indeed, in the published paper [5] submitted to the IEEE Trans. on Image Processing, *in July of 1994,* we also recommended a similar scheme and method.

Conclusions

We have shown that the two main parts of both papers are practically identical. Our paper [1] is included among the references of [2], but is cited only once, and as a reference to the S-Transform—which can be found in several other places. We have shown that the new proposals are practically equal in their basic part, and that there is a large number of similarities throughout the papers. It was reported to us by several researchers that just glanced at both papers, and thus it is inconceivable that the CREW authors did not notice the resemblance. This proves that the CREW authors knew about our work, and did not give us credit because they did not want to.

Maybe it is worth adding that our lossless compression results were better, as shown in the table below. The CREW lossy compression results were presented in a way that makes fair comparisons practically impossible, but we strongly believe that our results were much better.

image	couple	lena
S+P [1]	4.08 bpp	4.25 bpp
CREW [2]	4.91 bpp	4.35 bpp

As mentioned above, improvements of the methods proposed in [1] are described in a paper [5] that was submitted to the IEEE Transactions on Image Processing in July of 1994, and has since been published in September of 1996. The codec implementing the method is publicly available at `ftp://ipl/rpi.edu/pub/EW_Code` or via the Internet site with URL `http://ipl.rpi.edu/SPIHT/`.

We know about the existence of a later CREW paper where the reference to our work was removed, but we do not think it is worth discussing the subject any further.

References

[1] A. Said and W.A. Pearlman, "Reversible Image compression via multiresolution representation and predictive coding," *Proc. SPIE Conf. Visual Communications and Image Processing '93*, Proc. SPIE 2094, pp. 664–674, Cambridge, MA, Nov. 1993.

[2] A. Zandi, J.D. Allen, E.L. Schwartz, and M. Boliek, "CREW: Compression with reversible embedded wavelets," *IEEE Data Compression Conference,* Snowbird, UT, March 1995.

[3] J.M. Shapiro, "Embedded image coding using zerotrees of wavelets coefficients," *IEEE Trans. Signal Processing,* vol. 41, pp. 3445–3462, Dec. 1993.

[4] A. Said and W.A. Pearlman, "Image compression using the spatial-orientation tree," *IEEE Int. Symp. on Circuits and Systems,* Chicago, IL, pp. 279–282, May 1993.

[5] A. Said and W. A. Pearlman, "An Image Multiresolution Representation for Lossless and Lossy Compression," *IEEE Transactions on Image Processing*, Vol. 5, pp. 1303–1310, Sept. 1996.