

**An Improved Transform Coder for  
Image Sequences Using Attributes  
of Difference Pictures**

by

**M.R. Civanlar  
and  
Pete Santago**

**Center for Communications and Signal Processing  
North Carolina State University**

**Raleigh, NC 27695**

**February 1985**

**CCSP-TR-85/2**

## ABSTRACT

One of the characteristics of the interframe difference pictures for video teleconferencing types of image sequences is large zero areas and impulse-like non-zero regions. It is possible to use this property to improve ordinary transform coders. In this paper, the method of projection onto convex sets is used to obtain an improved transform coder for impulse-like pictures in general and difference pictures in particular. Although the amount of improvement depends on many factors, including the percentage of motion, it is found significant for many cases.

## INTRODUCTION

Low bit rate image coding can be viewed as a restoration problem. In this sense, the receiver has incomplete information about the actual image and attempts to produce the most correct reproduction of it. Thus, a transform coded image which is specified by only a part of the transform coefficients of the original image is a degraded signal and can be modeled as:

$$f_t = (T^{-1} ST)f + n \quad (1)$$

where,  $f$  is the original image,  $f_t$  is the coded image,  $T$  is the transformation matrix,  $S$  is a modified identity matrix with zero diagonal elements corresponding to unused transform coefficients, and  $n$  is the noise caused by quantization.

If the only information available to the receiver is the transmitted transform coefficients, then  $f_t$  is the best estimate for the original image. Most of the time, however, this is not the case. The receiver may have knowledge about the statistics or other properties of the original image. This information may be free, as in the case of nonnegativity, or it may be sent as an alternative for some of the transform coefficients. A signal restoration algorithm can be used to improve the estimate for the original signal of eq. 1 by incorporating this additional information.

A suitable restoration technique for this application may be one using the method of successive projection onto convex sets (POCS) for finding points in the intersection of sets. The main idea of POCS presented in [1] is as follows: let  $C_1, \dots, C_m$  be closed and convex sets with nonempty intersection in a Hilbert space. The iterative sequence given by:

$$f^{k+1} = P_1[ P_2[ \dots P_m[f^k] ] ] \quad (2)$$

where,  $P_i[\circ]$  is the projection operator for the  $i^{\text{th}}$  set,) then converges to a point in the intersection for any initial estimate.

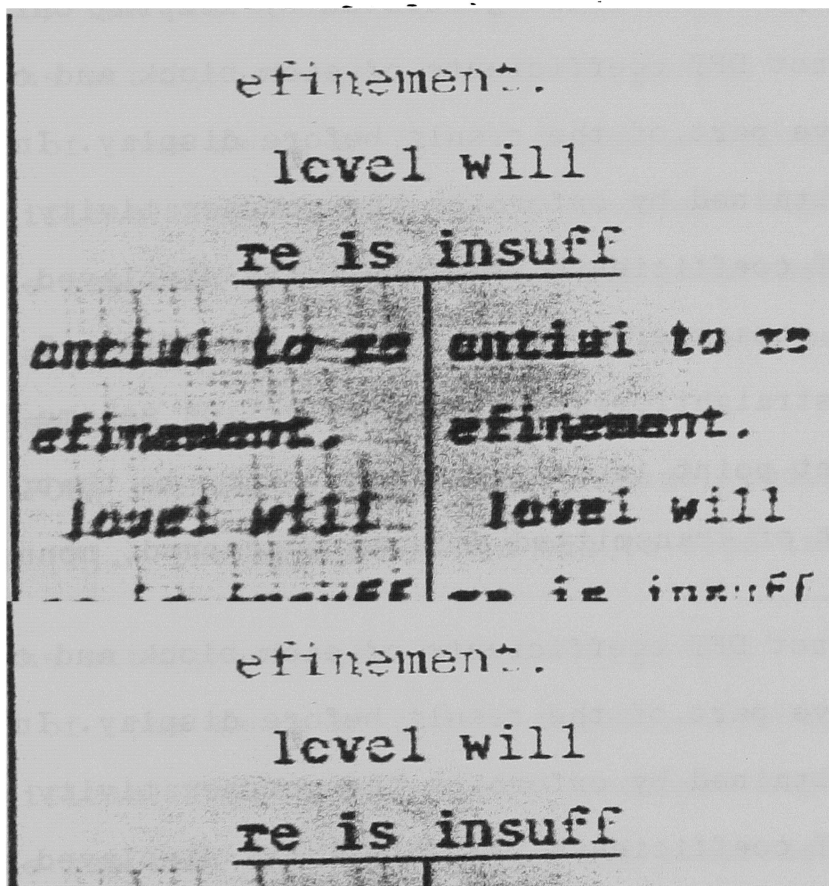
The set of signals defined by a certain number of transform coefficients is closed and convex. The projection of any signal onto this set can be obtained by making the transform coefficients of the signal equal to the specified ones. In this sense, transform coders use the projection of the zero signal onto this set as their estimate.

Any information which can be used to define a convex set in the signal space may be incorporated in the restoration algorithm to improve the transform coder. However, computational problems restrict the usable sets to those with simple projection operators. In this work the effect of two sets, those defined by nonnegativity and the region of support information will be demonstrated.

## NONNEGATIVITY

The effectiveness of nonnegativity in the restoration of impulse-like signals has been reported earlier [2],[3]. Nonnegativity of the intensity is a common property of all physical images and it can be used to enhance transform coders. However, in order to achieve a noticeable improvement by incorporating this information, the coded image must have regions with zero intensity. Certain classes of images, such as text, actually have this property, and some images may easily be modified by subtracting a bias so that they have large zero regions.

In order to demonstrate the effect of nonnegativity, an actual 128X128 text image is transform coded by using 32X32 blocks and the DFT. In fig. 1.a the original image is



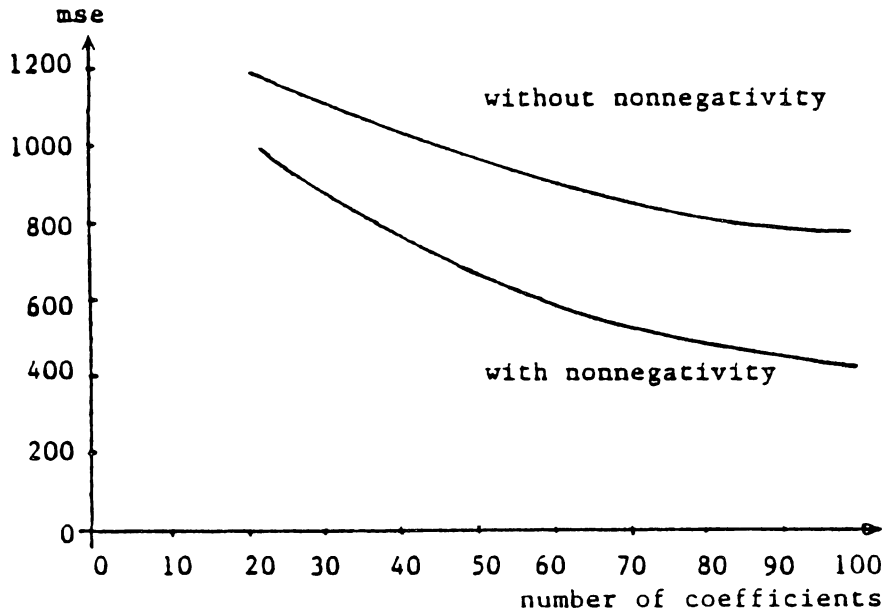


Figure 2. Improvement due to nonnegativity

given. It should be noted that the original image is not a binary signal but has intensity variation from 0 to 255. Fig. 1.b displays the result generated by keeping only 45 out of 512 distinct DFT coefficients of each block and clipping the negative part of the result before display. In fig 1.c the result obtained by enforcing the nonnegativity and the specified DFT coefficients iteratively is displayed. The reduction in the square error is displayed in fig. 2 for determined by a straightforward zigzag selection scheme [4].

The important point in using nonnegativity is that, as far as the amount of transmitted data is concerned, nonnegativity is free information. The price of improvement is the required iterations. With special purpose hardware, however, these iterations can be carried out easily.

## CODING THE DIFFERENCE PICTURE

There are various coding techniques using the inter-frame difference pictures of video teleconferencing image sequences [5],[6]. In hybrid coding, successful results in low bit rates are obtained through the use of transform coding and DPCM combination. Motion compensation can also be used to further improve these results [5]. The technique using the POCS method for improving the transform coding can also be used in combination with these methods. That is, the transform coefficients can be transmitted using DPCM in space or time domain and/or motion compensation can be employed for improving the quality at lower transmission rates. However, since the scope of this paper does not cover the design of a complete coder, the improvement obtained by the new method will be demonstrated by comparing it with the straightforward transform coding. The incorporation of the above mentioned techniques would obviously give significant improvement to the proposed algorithm, and thus, the reported quality figures and compression rates do not reflect the typical performance of the new method in a complete scheme.

The POCS technique for general image coding is discussed in [7]. The approach used was to send information to describe convex sets of images which include the frame to be coded and to find an element in the intersection of these sets by using the POCS method. If the intersection is small enough, any member of it will be a good approximation for

the original frame. Various sets and their interactions are demonstrated in [7]. Clearly, certain groups of images have attributes which can be used for defining sets which effectively restrict the size of the intersection set. The interframe difference pictures of video teleconferencing image sequences do have such a property, namely, large zero regions.

A typical interframe difference picture is displayed in fig. 3.a. (The original images are 128X128 with intensity variation from 0 to 255.) The combination of nonnegativity with the transform coding seems to be suitable for this type of image. Unfortunately, difference pictures are not nonnegative. It is possible to obtain some improvement by coding the nonnegative and nonpositive parts of the difference picture separately in order to invoke nonnegativity. However, since the number of frequency coefficients used for each part have to be half of the number used in coding the difference picture directly, improvement is not significant, especially for high compression rates.

The technique used in this paper to take advantage of this attribute is to transmit the zero/nonzero information by sending one bit per pixel. This is equivalent to the region of support information used in bandlimited signal extrapolation [8]. Once the location of the nonzero pixels are available, another bit per nonzero pixel can be transmitted for the sign and the picture can be treated as nonnegative. The number of transform coefficients is



reduced to compensate for the additional bits needed to transmit the pixel domain information described above. Certainly, a suitable coding technique, such as a run length code, can be used to convey this pixel domain information using less than one bit per pixel. Also, a subsampling technique can be used as in [7]. In our comparisons with difference picture transform coding, however, these reductions are not considered.

The effectiveness of the region of support and sign information is a function of the number of nonzero pixels and sign changes. In some of the cases, sending extra frequency coefficients may be more beneficial than sending the stated space domain information. In order to take advantage of this fact each block is coded using the best combination. In table 1, the number of blocks coded with smallest error using each combination of transform coefficients, zero locations, and non-zero pixel sign, is presented for different block sizes. In fig 3.b and 3.c the results obtained by coding the difference picture shown in fig 3.a by using an ordinary transform coder and the new technique are demonstrated respectively. For this picture a blocksize of 16x16 and a 4:1 compression ratio is used. The reduction in mean square error is 56%. In table 2 the mean square error and the improvement over the straightforward transform coder are displayed for different block sizes and amounts of motion.

There is an obvious improvement as can be seen from the tables. As stated above, the ultimate compression rate can be made many times higher than those stated in the tables by incorporating different coding techniques in the actual transmission.

Table 1

<u>Blocksize</u>	<u>Trans</u>	<u>Loc &amp; Trans</u>	<u>Loc, Sign &amp; Trans</u>
1 x 128	45	37	46
8 x 16	7	64	57
8 x 32	4	20	40
16 x 16	0	31	33

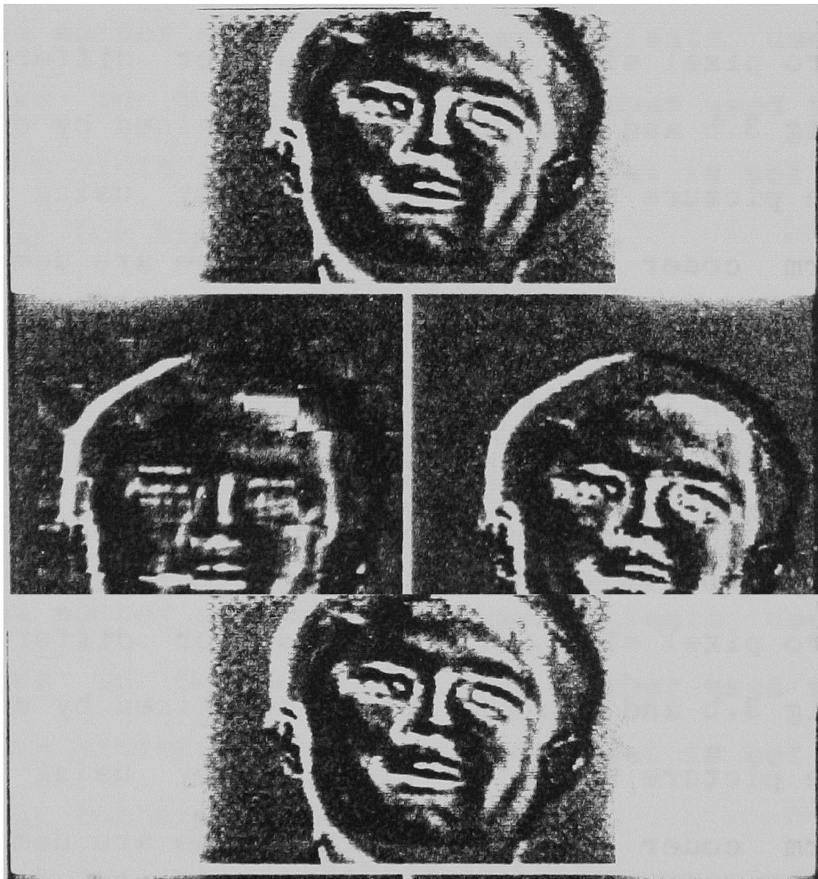
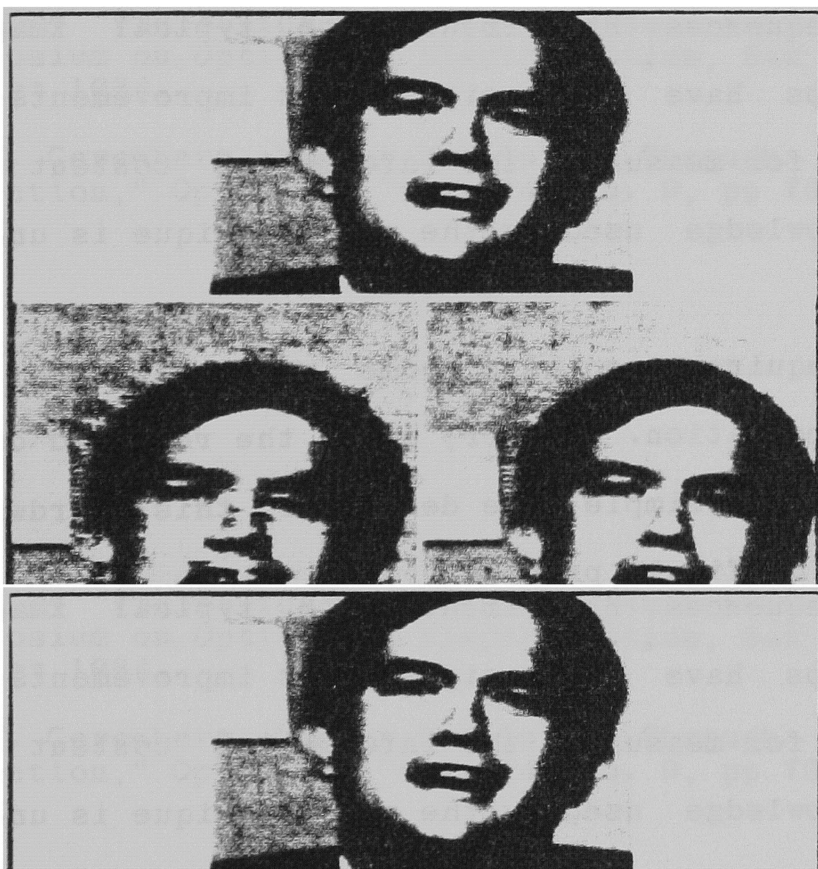


Table 2

<u>Blocksize</u>	<u>15% motion</u>		<u>7% motion</u>	
	<u>mse</u>	<u>improve</u>	<u>mse</u>	<u>improve</u>
1 x 128	11.8	17%	1.6	63%
8 x 16	9.3	41%	1.9	57%
8 x 32	9.5	45%	1.8	62%
16 x 16	9.4	56%	2.3	59%

As in all other schemes using the difference pictures, this method suffers from error accumulation. The result obtained in the fourth frame of the sequence is displayed in fig. 4. The quality of the picture obtained using the new technique is still acceptable while the transform coder has degraded the image much more severely. In figure 5 the square error versus the frame number is displayed for both. Clearly, an update mechanism is necessary and will be effective for determining the final compression rate.



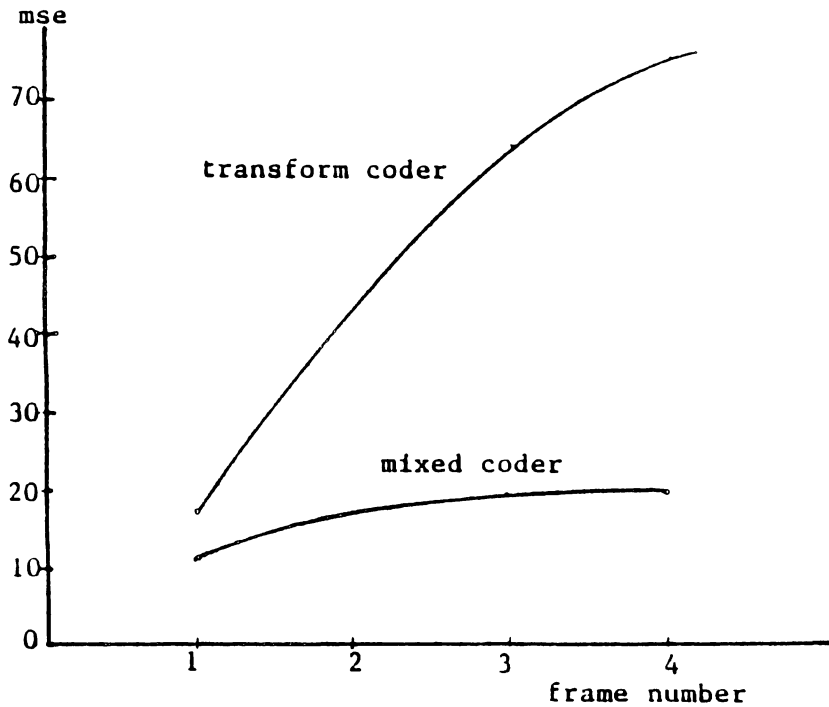


Figure 5. Error Accumulation

#### SUMMARY

Using the POCS technique, it is possible to improve the transform coding of impulse-like images and video teleconferencing image sequences. Experiments with typical images from these groups have shown significant improvements. A theoretical study for measuring the information content of the combined knowledge used by the new technique is under investigation.

The method requires special purpose hardware for the real time implementation. However, since the required data manipulations are very simple, the design of this hardware should not pose a difficult problem.

## REFERENCES

- [1] L.M. Bregman, "The Method of Successive Projection for Finding a Common Point of Convex Sets," Doklady, Akad. Nauk. SSSR, vol. 162, no. 3, pp. 688-692, 1965.
- [2] R.N. Schafer, R.M. Mersereau, M.A. Richards, "Constrained Iterative Restoration Algorithms," Proc. of the IEEE, vol. 69, no. 4, April 1981.
- [3] H.J. Trussell, M.R. Civanlar, "Feasible Solution in Signal Restoration," IEEE, Trans. Acoustics Speech and Signal Processing, vol. ASSP-32, April 1984.
- [4] J.A. Szarek, "Hybrid Coding Image Bandwidth Reduction Algorithm," M.S. Thesis, North Carolina State University, Raleigh, NC, 1983.
- [5] T.S. Huang (editor), Image Sequence Analysis, Springer-Verlag, New York, 1981.
- [6] W.A. Pearlman and P. Jakatdar, "The Effectiveness and Efficiency of Hybrid Transform/DPCM Interframe Image Coding," IEEE Trans. on Comm., vol. Com-32, No. 7, pp. 832-838, July 1984.
- [7] P. Santago, S.A. Rajala, "Techniques for Video Compression by Projection onto Convex Sets," SPIE 28th Annual Symposium on Optics and Electro-Optics, San Diego, CA, August 1984.
- [8] R.W. Gerchberg, "Super Resolution Through Error Energy Reduction," Opt. Acta, vol. 14, no. 9, pp 709-720, Sept. 1979.