December 4, 1984

Technique For Video Compression By Projection Onto Convex Sets

P. Santago, *North Carolina State University*
Sarah A. Rajala, *North Carolina State University*

Technique for video compression by projection onto convex sets

P. Santago and S.A. Rajala

Center for Communications and Signal Processing
North Carolina State University
Department of Electrical and Computer Engineering
Raleigh, North Carolina

Abstract

This paper describes a video compression technique which utilizes the alternating projection theorem for convex sets. The image to be transmitted is determined to be in certain convex sets and parameters defining these sets are sent. The receiver can then use the method of successive projections to locate an image which is in the intersection of the sets. If the intersection is small then the image determined should be close to the desired image. The coder can be made more robust by easily adding additional convex sets or using it in conjunction with other coding schemes such as motion compensation.

Introduction

Video bandwidth compression has received much attention in recent years as is evidenced by the research accomplished in the area.[1] Techniques in spatial as well as transform domains have been utilized to achieve compression. Using unitary transforms, it is hoped that a few coefficients will adequately define the desired image. Further, in motion sequences, an attempt to predict frame-to-frame movement may be used to reduce bandwidth.

The theory of convex sets and the alternating projection theorem[2] make it possible to define image properties much like the idea of using transform coefficients. Parameters of certain convex sets can be transmitted, for example the energy of the desired image, and an image exhibiting the transmitted properties can be found using alternating projections.

Convex Sets

A convex set, C, is a set such that given any two elements in C, all points on the line segment connecting the two elements are in C. More formally, given a set C in a linear vector space, C is convex if given any $x,y \in C$, then $[\alpha x + (1-\alpha)y]$ is in C for all $0 \leq \alpha \leq 1$ (see Figure 1).

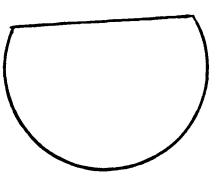
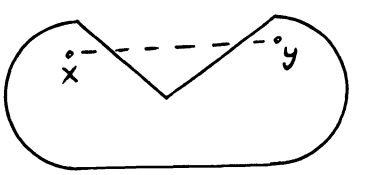


Figure 1(a). Convex



(b). Not convex

Two theorems from convex set theory are stated without proof.

Theorem 1[3]
Let x be a vector in a Hilbert space H and let K be a closed convex subset of H. Then there exists a unique $k_0 \in K$ such that

$$\|x - k_0\| <= \|x - k\| \text{ for all } k \in K$$

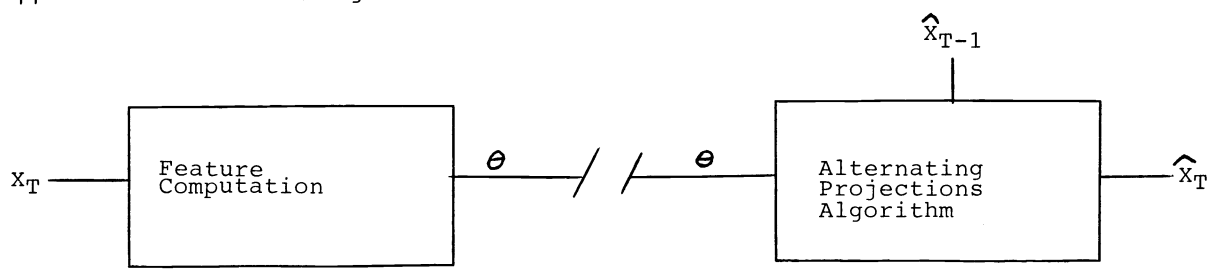
The vector k_0 is the projection of x onto the set K, indicated by $P_K(x)$.

Theorem 2[2]
Let $I = \bigcap_{i=1}^{n} S_i$ be nonempty and S_i be a closed convex set in a finite dimensional Hilbert space. Then the iterative procedure

$$\hat{x}_{(k+1)} = P_1 P_2 \ldots P_n(\hat{x}_k) \tag{1}$$

converges to a point in I.

Theorem 2 gives a procedure for locating a point in the intersection of a collection of closed convex sets if the individual projections are known. Further, if the intersection is of small size or of a certain shape, the fixed point found by (1) may be a good reproduction of the desired image. Put another way, given an initial estimate, x_0, and features of the original image, x, as defined by some collection of convex sets, force x_0 to have the same features by projecting x_0 to the intersection of the sets. A coder using this approach is shown in Figure 2.



Θ - vector of set parameters

Figure 2. Convex Set Projection Coder.

Set selection

Choosing convex sets that will give the best results in the sense of bandwidth reduction or image reproduction is not an easily solved problem in the most general case. Intelligent choices may be made, however, using some results from image restoration.[4,5,6]

In Curtis' paper on reconstruction from one bit of Fourier phase[5], it can be seen that the phase contains much information about the image. Also, all images whose phases agree in sign form a closed convex set. Examining Civanlar's work[6], it may be possible to model the noise due to motion in time varying image sequences to aid in image reproduction. This approach is desireable also since the convex sets involved may not necessarily involve transmission, much like backward motion compensation.[7]

Other possibilities may be gotten by combining information in the spatial and frequency domains. An example of some convex sets of this type would be to indicate a collection of correct values in time and the sign of the real part in frequency (see (2) and (3)) and project to the intersection which satisfies these values in both domains.

$$\{x \varepsilon R^n \mid x_i = k \text{ ,k a constant, i given}\} \tag{2}$$

$$\{x \varepsilon C^n \mid \text{sign } Re[x_i] = + \text{ or } - \text{ is given, i given}) \tag{3}$$

where x_i is the i'th element of x.

In the following sections are three examples that use the technique described for coding and image in a time varying sequence. Example 1 shows how conex sets can be employed with an existing coding method. Example 2 uses the two sets defined by (2) and (3). Finally, example 3 illustrates the effect of the initial estimate on the image which the algorithm locates.

Example 1

This example uses a head scene of a newscaster who is speaking and moving his head. The image is 32x256 pixels and the moving area is about 32x120 pixels. In a typical transform coding procedure, the scene is segmented into fixed size blocks, each block is transformed, and the significant coefficients are transmitted. In this example 4x8 blocks were used and the first nine coefficients (2 real and 7 complex) were kept using a zigzag priority assignment.[8,9] Coefficients not transmitted were assumed to be zero resulting in a low pass image when the transformed block is inverted.

Using projections, the previous image is transformed in the same manner and the transmitted coefficients are put into corresponding positions in the transformed blocks keeping the untransmitted coefficients unchanged. This is the projection onto a convex set as described in (2) with C^n instead of R^n. Table 1 shows the error statistics for the moving region. It should be noted at this time that coefficients were transmitted for either technique only if the low pass technique would benefit from transmission, that is only in regions with significant change.

Questions remain as to how well this process performs when the high frequency coefficients are kept from frame-to-frame. It would probably prove necessary to trade off some improved quality for bandwidth to update the high frequency coefficients periodically. In any event, if high frequency components are not changed, information about the quality of the resultant image may be derived by examining the intersection of convex sets defined by low frequency basis vectors along with information about the initial image used in (1).

Table 1.

	Low pass	Projection
# errors with magnitude > 0	3141	3071
# errors with magnitude > 3	1063	604
Maximum error magnitude	28	17
Square error	90551	29887
Signal-to-noise ratio	34.44dB	39.25dB

Example 2

The scene in this example is the same used in example 1. In the moving region there are 120 4x8 blocks. In example 1, to transmit the nine coefficients, 16 real numbers must be coded, say with four bits each on average resulting in 7680 bits being transmitted. In this example, the convex set defined by (2) is used by sending a 16 bit pattern indicating the correctness or incorrectness of every other pixel in the time domain. Another 16 bit pattern is sent to indicate the sign of the real part of the Fourier coefficients as in (3). Two additional convex sets are sent which indicate the mean pixel value and the inner product of the desired image with the difference picture of the previous two frame estimates (see (4) and (5)).

$$\{x \varepsilon R^n \mid \frac{1}{n} \sum_{i=1}^{n} x_i = m, \text{ m a constant}\} \tag{4}$$

$$\{x \varepsilon R^n \mid \langle x,d \rangle = k, \text{ k a constant}\} \tag{5}$$

Assuming 8 bits each to transmit m and k, then this technique requires the transmission of 5760 bits. Error statistics are given in Table 2.

Table 2.

	Before projection	After projection
# errors with magnitude > 0	3507	3356
# errors with magnitude > 3	2068	1199
Maximum error magnitude	42	28
Square error	324968	98356
Signal-to-noise ratio	28.89dB	34.08dB

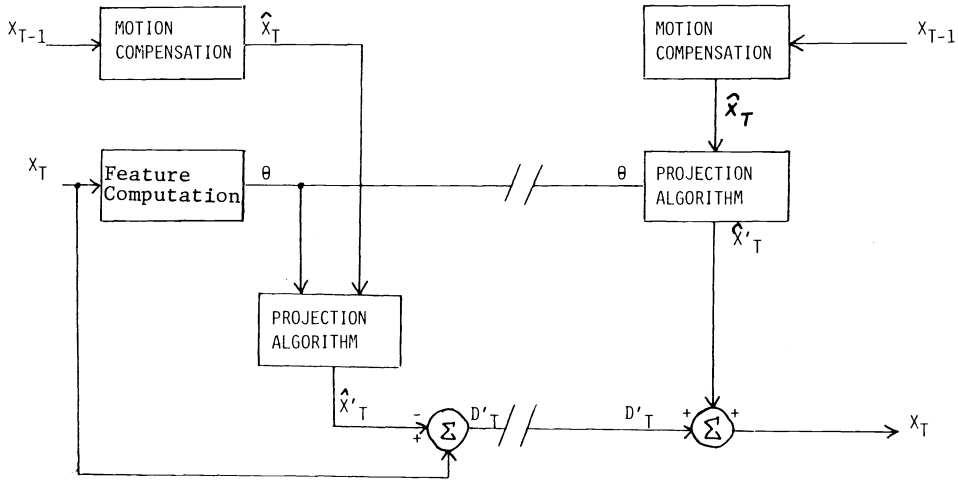
It should be noted in this example that sets such as (4) and (5) are merely basis function coefficients. However, linear bases are convex and therefore are easily incorporated into the procedure. Also, all three examples utilized a "clipping" set which requires that pixel values be in the interval [0,255]. This set requires no transmission, however. If a much smaller pixel range is known for a particular image, it may be advantageous to transmit this range to form a smaller convex set for clipping.

Example 3

This example is identical to example 2 except that the starting estimate is gotten from a backward motion compensation algorithm which needs no transmission[7] (see Figure 3). Also, the inner product set was not used which reduces even more the bandwidth required. Results are given in Table 3.

Table 3.

	Motion compensated	Projection
# errors with magnitude > 0	3084	3079
# errors with magnitude > 3	548	379
Maximum error magnitude	42	27
Square error	86110	44708
Signal-to-noise ratio	34.66dB	37.50dB



Θ - Feature set

Figure 3. Feature compensation coder with motion compensation

Discussion

As illustrated by the examples and from other simulations, it can be seen that the projection technique results in information compression. At this time it is not clear as to how much compression or what the limits of the method are, in particular as compared to coding techniques. The fact that projections allow non-linear relations to be used on images at least gives the potential for improvement.

Initial indications are that projection methods perform well when used with other techniques such as motion compensation as shown in example 3. In this case a better starting estimate is provided which should result in locating a point in the intersection which is "closer" to the desired image. Although this fact cannot be shown in general, simulation seems to bear it out, and so, in time varying images, the projection can make good use of frame-to-frame redundancy.

Summary

An image coder which transmits convex set parameters and uses the alternating projection algorithm for decoding is presented. The coder is shown to give bandwidth reduction using a newscaster scene. The coder is easily expanded by adding the projections for any new convex sets.

It should be noted, however, that the research on the use of convex sets in video compression is not complete. Problems which we are addressing in continued research include: convex sets selection, computational complexity, algorithm convergence rates, and the relationship between the intersection size and shape and the order or projection to image quality.

References

1. Huang, T.S. (ed). Image Sequence Analysis. Springer-Verlag, Berlin, 1981.
2. Bregman, L.M. "Finding the Common Point of Convex Sets by the Method of Successive Projections", DoKL. Akad. Nauk SSSR, 162, 3, pp.482-490, 1965.
3. Luenberger, David G. Optimization by Vector Space Methods. John Wiley & Sons, N.Y., 1969.
4. Youla, D.C. & Webb, H. "Image Restoration by the Method of Convex Projections", IEEE Trans. on Medical Imaging, Vol. MI-1, No. 2, October 1982.
5. Curtis, S.R., Lim, J.S., & Oppenheim, A.V. "Signal Reconstruction from One Bit of Fourier Transform Phase", Proceedings ICASSP 84, pp.12A.5.1-12A.5.4.
6. Civanlar, M.R. Digital Signal Restoration Using Projection and Fuzzy Set Techniques. PhD Dissertation, North Carolina State University, Raleigh, N.C., 1984.
7. Ishiguro, T. & Iinuma, K. "Television Bandwidth Compression by Motion-Compensated Interframe Coding", IEEE Comm. Magazine, Nov. 1982, pp.24-30.
8. Chen, Wen-Hsiung & Pratt, W.K. "Scene Adaptive Coder", IEEE Trans. on Comm., Vol. Com-32, No. 3, pp.225-232.
9. Szarek, J. A Hybrid Image Bandwidth Reduction Algorithm. M.S. Thesis, North Carolina State Univ., Raleigh, N.C., 1983.