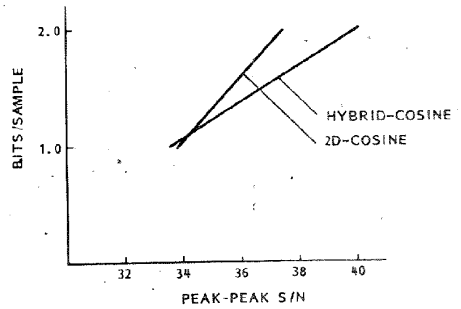


Fig. 7. Comparison of adaptive hybrid and 2-D cosine transforms.

sional transform coding technique has a better performance at lower bit rates.

REFERENCES

[1] A. Habibi, "Survey of adaptive image coding techniques," *IEEE Trans. Commun.*, vol. COM-25, pp. 1257–1284, Nov. 1977.
[2] A. G. Tescher and R. V. Cox, "Image coding: Variable rate DPIM through fixed rate channel," in *SIPE Proc.*, vol. 119, San Diego, CA, Aug. 1977, pp. 147–154.
[3] R. F. Rice and J. R. Plaunt, "Adaptive variable-length coding for efficient compression of spacecraft television data," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 889–897, Dec. 1971.

★

Ali Habibi (S'69–M'70–SM'76), for a photograph and biography, see this issue, p. 1727.

Interfield Hybrid Coding of Component Color Television Signals

FARHAD A. KAMANGAR, MEMBER, IEEE, AND K. R. RAO, SENIOR MEMBER, IEEE

Abstract—**Hybrid coding of color television images at broadcast standards is investigated. This involves two-dimensional discrete cosine transform (2-D DCT) of each field, followed by the differential pulse code modulation (DPCM) between successive fields. Two 22-frame sequences of "Water Skier" and "Wheel of Fortune" in component form, i.e., luminance Y and chrominance I and Q, were utilized as the database. A statistical study of the prediction error in the 2-D DCT domain is carried out. Three different algorithms for the 2-D DCT system are simulated and analyzed. In the first algorithm, optimum (nonuniform) quantizers are used in DPCM loops followed by fixed-wordlength coders. Uniform quantizers along with variable-wordlength coders are implemented in the second system. The performances of different deterministic coders are investigated and compared with Huffman coders. The third scheme is an adaptive system. In this system each 8 × 8 block is divided into four subblocks. The activity of each subblock is monitored and when it exceeds some threshold, the subblock is considered to be spatially active. More bits are assigned to active subblocks and the ranges of corresponding quantizers are expanded. The prediction error of the dc coefficient is monitored to determine the temporal activity of a block. For a temporally active block, the number of bits assigned to the two lower frequency subblocks is increased. Performance of these three systems for unmatched statistics and in the case of a scene change are studied. Mean square error (MSE) between the original and reconstructed images, bit rate, entropy, essential maximum, and other parameters are utilized as the performance criteria. Also, a subjective evaluation of the processed images is carried out. It is shown that the interfield adaptive hybrid coding of color TV signals in component form results in significant savings in bit rates for transmission over a digital link.**

Manuscript received January 13, 1981; revised September 28, 1981. This paper was presented in part at the National Telecommunications Conferences, Washington, DC, 1979, and Houston, TX, 1980, and is based on the research by F. A. Kamangar as a partial requirement for the Ph.D. degree from the University of Texas at Arlington, Arlington, TX. This work was supported by the Commercial Telecommunications Group, Rockwell International, Dallas, TX.

The authors are with the Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX 76019.

INTRODUCTION

DIGITAL transmission of images has gained considerable importance [1]–[20], [42]–[45] in view of its application to satellite and carrier communication of television images (picturephone, conference, classroom, industrial and network TV—both monochrome and color). Major research in this field is focused on bandwidth (BW) compression, i.e., removing the redundancy inherent in an image or sequence of images both in space and time such that the images can be transmitted at reduced bit rates. For network quality color TV digital transmission, the fidelity requirements are very rigid [2], [3], [7]–[11], [21]. Yet substantial BW reduction can be achieved by considering the image redundancy and psychovisual characteristics of the human vision. Many data compression techniques for picturephone, conference TV, industrial TV, satellite images from outer space, and images from remotely piloted vehicles have been proposed and implemented. Also, prototype commercial systems based on predictive or transform coding have been designed, built, and are being marketed [12], [22]–[24]. Both intraframe and/or interframe coding applied to color images in composite or component form have been investigated. Initially the redundancy reduction has been attempted by predictive coding.

0090-6778/81/1200-1740$00.75

This was followed later with transform techniques as fast algorithms and special processors were developed. Combination of the two schemes (prediction and transform) also has been suggested and applied. This combination, called hybrid coding, acts as a compromise of terms of performance, capabilities, complexity, and limitations [25].

OBJECTIVE

The object of this paper is to develop hybrid [25] (2-D DCT/DPCM) coding of color TV (component form) for digital transmission at reduced bit rates. Data compression is based on 2-D DCT [26], [27] of each field followed by DPCM [28] between successive fields. Based on the histogram of the predictor error in the 2-D DCT domain, quantizers for minimizing the mean square quantization error (MSQE) [30] are developed. The performance of these optimal (nonuniform) quantizers with fixed length coders is compared with the uniform quantizers coupled with variable length coders. Performance criteria such as mean of the absolute reconstruction error (MARE), absolute maximum reconstruction error (AMRE), bit rate, entropy, and essential maximum [29] are utilized for evaluating these quantizers.

DATABASE

The database for the hybrid simulation consists of 22 frame sequences each of "Water Skier" and "Wheel of Fortune" in component form (Y, I, Q), i.e., luminance Y plus chrominance I and Q. The amplitude of the picture elements (pels or pixels) has been originally quantized uniformly to six bits (64 levels). This is increased to eight bits by adding two zeros as the two least significant bits to the six-bit quantizer. Each of the Y, I, and Q components is sampled at 8.064 MHz with 416 pels/horizontal line and 464 lines/frame (visible portion). The database was supplied in this format by NASA Ames Research Center (ARC). The bandwidths of Y, I, and Q are 4.2 MHz, 1.5 MHz, and 0.5 MHz, respectively.

HYBRID CODING

Introduction

The concept of hybrid coding for video signal is to remove the inherent redundancy in the picture by a unitary transform in one or two dimensions followed by DPCM along the other dimension. In this system [25] each picture is divided into smaller blocks. Each block is passed through a 2-D transform and then a bank of DPCM loops removes the redundancy between the corresponding transform coefficients in the consecutive frames (fields).

Block Size

The choice of the block size is dependent on two factors. 1) The larger block size decorrelates more samples which will result in higher compression ratio. 2) The larger block size requires more arithmetic operations for forward and inverse transforms. It also requires more DPCM loops which will result in increased complexity in the final hardware realization. The block size was chosen to be 8 × 8 pels based on these considerations.

Interframe/Interfield Processing

Because two fields are interlaced to form one frame, the question arises if the blocks should be formed by adjacent lines in a field or in a frame. The factors that affect this choice are as follows. 1) Based on the 4 to 3 aspect ratio and the sampling rate of 8.064 MHz, the corresponding pels in the adjacent lines of a frame have higher correlation compared to the adjacent pels in each line. Also, corresponding pels in the adjacent lines of a field have lower correlation than the adjacent pels in a line. Thus, if the blocks are chosen in each frame, rather than in each field, it will result in higher decorrelation in the transform domain. 2) The interfield process will result in more correlation in the temporal direction, which results in higher efficiency of DPCM loops. This compensates for the lower decorrelation by the forward transform. 3) The interframe process increases the memory requirements in two parts of the hybrid system. a) For the transform process, since the lines are chosen in each frame, the first field should be held in a memory before the complete block can be formed and processed. This requires one field to be stored in both the transmitter and receiver, before applying the forward and inverse transforms. b) The shift registers used in the DPCM loops as the predictors should be doubled in size because, in this case, the differences between the corresponding DCT coefficients in successive frames are formed and quantized in these loops. Based on these considerations and simulation of the hybrid system for interfield and interframe processes, the interfield process was chosen for further study.

2-D DCT/DPCM

The array of elements resulting from 2-D DCT [26], [27] can be expressed as

$$y(u, v, k) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x(i, j, k) D(u, v, i, j); \quad u, v = 0, 1, \cdots, N-1 \tag{1}$$

where $x(i, j, k)$ and $y(u, v, k)$ represent the $(N \times N)$ arrays of the kth field in data and DCT domains, respectively, and $D(u, v, i, j)$ is the 2-D DCT kernel which is separable. The prediction error for each transform coefficient is defined as

$$e(u, v, k) = y(u, v, k) - \hat{y}(u, v, k); \quad u, v = 0, 1, \cdots, N-1 \tag{2}$$

where $\hat{y}(u, v, k)$, the predicted value of $y(u, v, k)$, is given by

$$\hat{y}(u, v, k) = \hat{y}(u, v, k \quad 1) + e_q(u, v, k-1); \quad u, v = 0, 1, \cdots, N-1. \tag{3}$$

$e_q(u, v, k-1)$, the quantized value of $e(u, v, k-1)$, is transmitted through the channel after coding. The image is reconstructed at the receiver side by the following inverse 2-D DCT/

DPCM operations:

$$\tilde{y}(u, v, k) = \tilde{e}_q(u, v, k) + \tilde{y}(u, v, k-1); \quad u, v = 0, 1, 2, \cdots, N-1 \tag{4}$$

$$\tilde{x}(i, j, k) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \tilde{y}(u, v, k) D^{-1}(u, v, i, j); \quad i, j = 0, 1, \cdots, N-1$$

where $D^{-1}(u, v, i, j)$ is the 2-D IDCT kernel. The difference $x(i, j, k) - \tilde{x}(i, j, k)$ is defined as the reconstruction error.

Statistical Model

In order to find the quanitzer characteristics and the number of quantization levels assigned to the prediction error of each DCT coefficient, the *a priori* knowledge of its statistics is required. To find the statistical parameters, such as mean and variance of each prediction error, a statistical model for the original picture should be established. It is assumed that each 8 × 8 block of pels represents a two-dimensional separable wide-sense stationary process. The statistics of the pels along horizontal and vertical directions are governed by an I-order Markov process. Based on this, the variance distribution of the prediction error can be easily determined.

Covariance Matrices

The correlations between the adjacent pels in each line and between the corresponding pels in the adjacent lines of a field were found to be 0.94 and 0.92 for Y, 0.97 and 0.96 for I, and 0.98 and 0.97 for Q, respectively. The theoretical variance matrix (based on an I-order Markov process) for the rows of Y in the DCT domain is shown in Table I. The variances of DCT coefficients based on analytical and experimental models are shown in Table II for Y only.

In order to find the statistics of the prediction errors, a hybrid system with no quantizers in DPCM loops is simulated. It is assumed that the presence of the quantizers will not have considerable effect on these statistics. The variance matrices of prediction errors for Y only are shown in Table II. For the theoretical model (I-order Markov process) temporal correlations were assumed to be 0.95, 0.96, and 0.97 for Y, I, and Q respectively. The variance distributions for I and Q are similar to those of Y (see Table II).

Quantizers

Design of the quantizers for DPCM loops involves, in general, 1) finding the optimum allocation of bits assigned to each prediction error for minimum mean square quantization error (MSQE), assuming that the total number of bits per block and the variances are known; and 2) determining the decision levels and output levels resulting in a minimum MSQE for a specific probability density function of the prediction error [30]. As the histograms of the prediction errors are approximately Laplacian, the optimum (nonuniform) quantizers were designed based on the Laplacian density function.

TABLE I
COVARIANCE MATRIX OF THE ROWS OF Y IN THE DCT DOMAIN

19223	1004	-285	80	-65	23	-20	5
1004	1590	-231	-16	-32	-4	-9	-1
-285	-231	845	-241	-15	-27	-3	-5
80	-16	-241	707	-243	-14	-24	-2
-65	-32	-15	-243	659	-242	-13	-18
23	-4	-27	-14	-242	639	-238	-11
-20	-9	-3	-24	-13	-238	630	-220
5	-1	-5	-2	-18	-11	-220	637

TABLE II
VARIANCE MATRICES OF THE PREDICTION ERRORS IN THE DCT DOMAIN Y. (a) THEORETICAL, (b) EXPERIMENTAL FOR WS SEQUENCE, (c) EXPERIMENTAL FOR WF SEQUENCE.

6534	541	287	240	224	217	214	217
742	614	33	27	25	25	24	25
399	33	18	15	14	13	13	13
335	28	15	12	11	11	11	11
313	26	14	12	11	10	10	10
303	25	13	11	10	10	10	10
300	25	13	11	10	10	10	10
304	25	13	11	10	10	10	10

(a)

6322	1631	439	91	25	12	24	31
2998	585	212	56	19	9	21	29
2145	401	166	50	18	8	20	27
1498	263	125	46	16	7	16	21
1113	201	94	40	13	7	11	14
871	153	79	34	12	6	9	11
607	121	65	29	10	6	6	7
523	111	55	27	12	6	16	26

(b)

2551	562	420	401	109	27	70	97
3026	672	496	259	75	25	61	95
3518	631	353	168	52	19	57	81
4918	743	250	123	39	15	50	74
4286	669	197	84	28	10	35	47
2188	477	132	63	21	7	18	27
930	284	103	52	18	6	8	9
506	197	87	44	21	6	21	34

(c)

Uniform Quantizers

In order to achieve lower bit rates, often variable length coding is applied to the quantizer outputs. It has been shown [31] that for a large number of output levels the entropy of the equally spaced (uniform) quantizer is less than or equal to the entropy of the optimum (nonuniform) quantizer, assuming that both quantizers have the same MSQE. This means that if the entropy coding is used, for a fixed bit rate output, the performance of the uniform quantizer in terms of MSQE is better than or equal to the nonuniform quantizer. Notice that the two quantizers being compared do not have the same number of output levels.

In terms of local visual distortions caused by the quantizers, the nonuniform quantizer has less granular noise because of its finer structure in the lower range. On the other hand, it also introduces more noise in the midrange input levels. Because, in the 2-D DCT/DPCM system, DPCM loops are removing the temporal correlation, the effect of using a uniform quantizer would be more granular noise on those parts of the picture which are not changing. To compare the performance of the nonuniform quantizer with uniform

quantizers, a set of uniform quantizers with optimum spacing between the output levels was designed assuming a Laplacian distributed input variable.

Bit Assignment

Using block coding technique, it is shown that for Laplacian distribution, the bit allocation for the prediction errors in the 8 × 8 block is governed by [36], [37]

$$\langle m_i \rangle = \frac{M}{S} + \frac{1}{2}\log_2 \sigma_{e_i}^2 - \frac{1}{2S}\sum_{j=1}^{S} \log_2 \sigma_{e_j}^2 \qquad (5)$$

where $\langle m \rangle$ stands for the nearest integer to m, M is the total number of bits assigned to the block, S is the number of samples in the block, and $\sigma_{e_i}^2$ is the variance of the ith prediction error. Modify (5) such that $\Sigma_{i=1}^{S} \langle m_i \rangle = M$. An example of the bit allocation for the prediction error for overall average rate of 2 bits/pel (BPP) is shown in Table III.

Coders

Variable length coding was used along with the uniform quantizers to remove further redundancy from the quantizers' outputs. Obviously, the most efficient coder would be the Huffman coder. The disadvantage of the Huffman coder is that it is very difficult to implement, especially for a large number of input messages, and it also needs *a priori* knowledge of the probability distribution of the messages. The other factor is that the efficiency of the Huffman coders decreases for unmatched statistics. For these reasons, other classes of deterministic coders were implemented and compared with the Huffman coders.

Laemmel Coders [32] of Length N, (L_N): In these coders the messages are divided into groups of length $2^N - 1$ messages each. Each message can be identified by its group number, $g = 0, 1, \cdots$, and the message number in its group $r = 1, 2, \cdots, 2^N - 1$. For each codeword the first gN bits are "0" and the last N bits are binary representations of r.

Golomb Coders [33] of Length N, (G_N): In these coders the messages are divided into groups of $2^{(N-1)}$ messages each. Each message can be identified by its group number $g = 0, 1, 2, \cdots$, and the message number in its group $r = 1, 2, \cdots, 2^{(N-1)}$. Each code is constructed by g bits of "0" as the most significant bits followed by bit "1." The next $(N - 1)$ bits are binary representations of r. It should be noted that G_1 and L_1 are the same.

PERFORMANCE CRITERIA

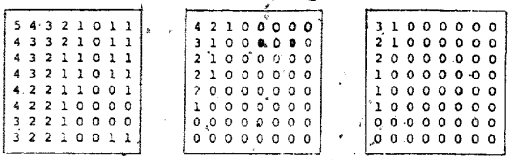
The performance of the hybrid coder can be judged by several criteria, both qualitative and quantitative. Some of the latter used to evaluate the system are as follows.

1) Mean of the absolute reconstruction error (MARE). MARE is defined as

$$\text{MARE} = E\{|x(i,j,k) - \hat{x}(i,j,k)|\} \qquad (6)$$

where $\hat{x}(i,j,k)$ is the reconstructed value of $x(i,j,k)$.

TABLE III
BIT ALLOCATION FOR THE PREDICTION ERRORS FOR OVERALL AVERAGE BIT RATE OF 2.0 BITS/PEL. (a) Y, (b) I, (c) Q.



(a) (b) (c)

2) Variance of the reconstruction error (ϵ^2). For a zero mean process the variance of the reconstruction error is defined as

$$\epsilon^2 = E\{[x(i,j,k) - \hat{x}(i,j,k)]^2\}. \qquad (7)$$

This is also the MSE between the original and reconstructed images.

3) Essential maximum (EM). The EM for Δ percent is defined such that the differences between the intensities of the original and reconstructed pels occur within EM, Δ percent of the time, or simply Δ percent of the reconstruction errors fall in the range of EM.

4) Absolute maximum reconstruction error (AMRE). AMRE for a process is defined as

$$\text{AMRE} = \max\left(|x(i,j,k) - \hat{x}(i,j,k)|\right). \qquad (8)$$

This shows the highest value of the reconstruction error caused by overloading the quantizers. AMRE also can be defined as 100 percent EM.

5) Signal-to-noise ratio in dB (SNR). The peak-to-peak signal-to-noise ratio (SNR) is defined as

$$\text{SNR} = -20\log_{10}\left(\frac{\sqrt{\epsilon^2}}{V_{pp}}\right) \qquad (9)$$

where V_{pp} is the peak-to-peak signal which for 8 bit PCM is equal to 255.

6) Entropy (H). The entropy of the quantized prediction error with N possible levels is defined as

$$H = -\sum_{i=0}^{N-1} p(q_i)\log_2\left(p(q_i)\right) \text{ bits} \qquad (10)$$

where $p(q_i)$ is the probability of the ith quantization level.

The ratio of $(\text{MARE})^2$ to ϵ^2 is a measure of granular noise of the system in comparison to the overload noise. For example, between two systems having equal variances of the reconstruction error, the one with larger MARE introduces more granular noise and less overload noise compared to the one with smaller MARE. A subjective evaluation of the original and reconstructed frame sequences yields the qualitative performance of the hybrid coder.

The statistics of the hybrid system for Y, I, and Q of the WS sequence are shown in Table IV. The lower AMRE of

TABLE IV
(a) STATISTICS OF THE HYBRID SYSTEM WITH NONUNIFORM QUANTIZERS AND FIXED LENGTH CODERS (Y OF WS SEQUENCE)
(b) STATISTICS OF THE HYBRID SYSTEM WITH UNIFORM QUANTIZERS AND FIXED LENGTH CODERS (Y OF WS SEQUENCE)

Average # of BBP	ϵ^2	MARE	AMRE	SNR in dB	EM 99%	EM 95%	Average entropy (BPP)
0.500	81.83	5.85	101	29.0	36	18	0.419
1.000	32.92	3.78	71	32.0	23	11	0.818
2.000	11.47	2.29	49	38.5	12	4	1.665
3.000	4.46	1.49	33	42.6	9	3	2.511

(a)

Average # of BPP	ϵ^2	MARE	AMRE	SNR in dB	EM 99%	EM 95%	Average entropy (BPP)
0.500	95.90	6.39	119	28.3	36	20	0.396
1.000	44.08	4.32	86	31.7	25	13	0.756
2.000	18.39	2.74	59	35.5	17	8	1.518
3.000	8.54	1.81	47	38.8	12	4	2.248

(b)

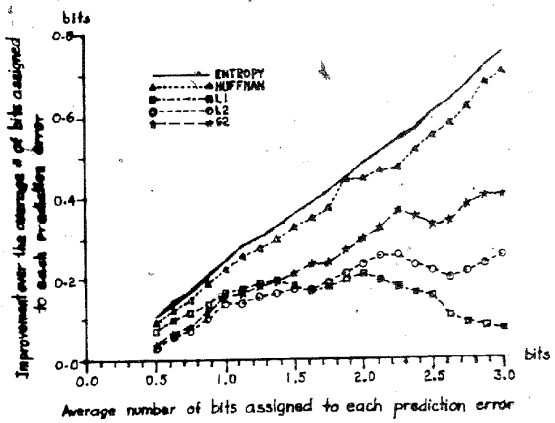


Fig. 1. Performance of different coders for uniform quantizers.

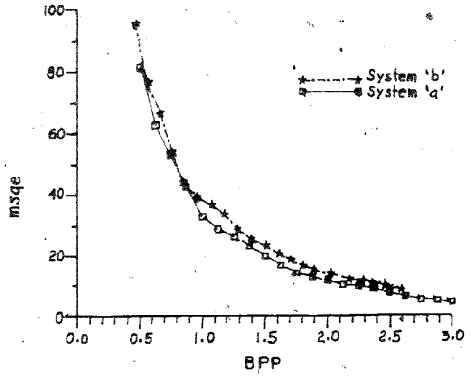


Fig. 2. MSQE versus average number of bits/pel for nonuniform quantizers and fixed length coders (system "a") and uniform quantizers with G_2 coders (system "b").

the system with nonuniform quantizers [Table IV(a)] compared to that for the uniform quantizers [Table IV(b)], for equal average bit rates, is the result of the wider range of the nonuniform quantizers. This wide range causes a faster recovery in temporal direction in the case of large frame-to-frame changes caused by a violent motion or a scene change.

Fig. 1 describes the performance of different coders used along with uniform quantizers. Since these coders do not have a constant output bit rate, a buffer is needed to smooth out the output bit rate. The average bit rates at the outputs of the L_3, G_3, L_4, and G_4 coders are most of the time larger than the average bit rates at the inputs to these coders [37]. This means that the use of these coders will decrease the efficiency of the system. It can be seen that for the average bit rates less than 1.375 bits/pel the L_1 coders should be used. Above this bit rate, G_2 coders result in lower output bit rate. Because the coders with longer wordlengths have better performance for a larger number of input messages, a combination of different coders can be used for different coefficients to improve the average bit rate of the system. Since the G_1 and L_1 coders are the same, we conclude that G_N coders have a better overall performance than L_N coders. Fig. 2 shows a comparison in terms of the MSQE between the hybrid system with nonuniform quantizers and fixed length coders (system "a") and the system with uniform quantizers and G_2 coders (system "b"). It can be seen that system "a" has lower MSQE at all different average bit rates. Since the number of bits assigned to the I and Q blocks is small, the variable length coders were not used for I and Q. The original and reconstructed images are shown in Figs. 3–8.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.