Statistical Performance Analysis of an Interframe Encoder for Broadcast Television Signals

TOSHIO KOGA, YUKIHIKO IIJIMA, KAZUMOTO IINUMA, AND TATSUO ISHIGURO

Abstract—This paper describes an objective evaluation for coding performance of an interframe encoder (NETEC-22H). Also described is the coding performance improvement by an adaptive bit sharing multiplexer (ABS-MUX) in which transmission bit rate is dynamically allocated to several channels.

Measurements made for actual broadcast TV programs over a time of 36 h show that an SNR of higher than 50 dB unweighted is obtained by this coding equipment for 99 percent of the time for broadcast TV programs at the transmission bit rate of 30 Mbits/s and for 93 percent of the time at 20 Mbits/s. The residual 1 percent at 30 Mbits/s or 7 percent at 20 Mbits/s is transmitted with a slightly lower SNR. The picture quality difference between the 20 and 30 Mbit/s transmission is about 6 dB in SNR on the average.

It is also shown that a three-channel ABS-MUX (20 Mbits/s per channel on the average) reduces probability of coarse quantization by a factor of 5–10 compared with the fixed bit rate transmission at 20 Mbits/s.

I. INTRODUCTION

VARIETIES of television coding algorithms have been devised and developed [1], [2]. Interframe coding has been expected to be most promising [3]. At first, it was applied to 1 MHz video telephone for face-to-face communication [4]. In due course, the application was extended to 4 MHz video teleconferencing. Interframe coding has paved the way to visual communication with full motion video [5]–[10].

Until now, emphasis has been placed on achieving a high compression ratio for video teleconferencing application. On the contrary, as far as broadcast TV signal transmission is concerned, it is more important to transmit a high quality signal than to achieve a high compression ratio. Some examples are mentioned in the literature of such digital television encoders for broadcast TV program transmission use at 16–30 Mbits/s [11]–[13].

The interframe encoder (NETEC-22H) described here is an improved version of that described in [11], in which an adaptive interframe/intraframe prediction is used to improve coding performance for pictures with substantial motion [14]. According to subjective evaluation, encoded picture quality is excellent for most pictures at a coding bit rate of 20–30 Mbits/s. However, if a very busy picture with active motion is supplied, some picture quality degradations are observed, although they are encountered with a very small probability in actual television programs. Thus, the interframe coding performance greatly depends upon picture content.

In order to obtain an objective measure of the interframe coding performance, knowledge of the statistics of TV signals is necessary. In this paper, probability distributions of the amount of information encoded by NETEC-22H [15] are measured for actual broadcast TV signals for many hours. Also, statistics of the encoder parameters, which are adaptively controlled according to the buffer memory occupancy or equivalently the rate of source signal information, are measured. These results are exploited in calculating SNR probabilities of encoded broadcast TV signals. Furthermore, effective utilization of transmission bit rate and picture quality improvement can be achieved by means of the proposed adaptive bit sharing multiplexer (ABS-MUX) for multiple simultaneous channel transmission [16], using the advantage of instantaneous differences among multiple channels which are statistically similar. The effect of ABS-MUX is also measured for actual broadcast TV signals by using physically realized hardware systems.

II. CODING ALGORITHM

The encoder/decoder block diagram is shown in Fig. 1. A composite NTSC color TV signal is normally sampled at 10.76 MHz and digitized into an 8 bit PCM signal, and then compressed to reduced bit rate data of 2–3 bits/picture element (bit/pel) on the average through digital signal processing. First, the preprocessor makes compensation for the phase inversion of the color subcarrier between two successive frames. The adaptive interframe predictive encoder removes redundancy from the signal by adaptive prediction and the variable wordlength coder compresses prediction error information with variable wordlength codes. The compressed data stream, which is generated irregularly, is smoothed out by the buffer memory and sent to the transmission line. At the decoder, the inverse processing is made to reproduce the NTSC color TV signal. The variable wordlength decoder expands the compressed data supplied from the buffer memory. The expanded data is decoded through the adaptive interframe predictive decoder to yield the phase-compensated signal. The postprocessor produces the composite NTSC color TV signal to be D/A converted.

Thus, the codec is designed so that the composite video signal is directly encoded and the waveform of the input signal is preserved except for quantization effect.

The encoder/decoder operates under two different sampling frequencies: 10.76 MHz ($684 \times f_H$, f_H: horizontal sync fre-

Manuscript received March 13, 1981; revised August 5, 1981. This paper was presented in part at the National Telecommunications Conference, Dallas, TX, December 1976, and the International Conference on Communications, Boston, MA, June 1979.

The authors are with Nippon Electric Company, Ltd., Kawasaki, Japan

0090-6778/81/1200-1868$00.75 © 1981 IEEE

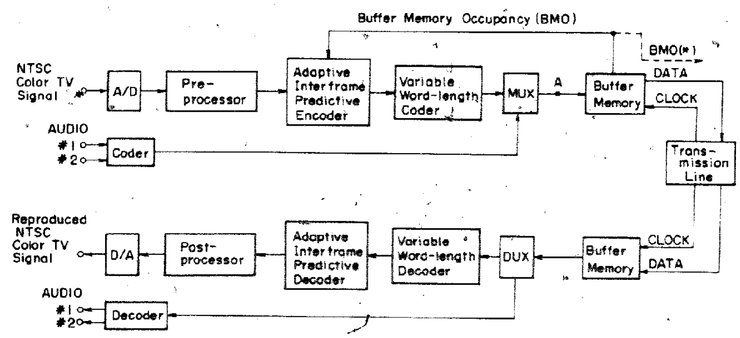

Fig. 1. NETEC-22H encoder/decoder block diagram. BMO (*) is supplied to ABS-MUX.

quency) in T-mode and 7.16 MHz ($456 \times f_H$) in S-mode. The T-mode is a normal operation mode. The S-mode is a sub-Nyquist frequency operation which is only applied for extreme cases where a large amount of information is generated and buffer fill occurs or is likely to occur. Therefore, T-mode operation is mainly described in what follows.

A. Subcarrier Phase Compensation

Since the subcarrier phase of the NTSC color TV signal alternates by 180° at a sampling point, frame by frame, subcarrier phase compensation is necessary before taking frame-to-frame differences.

Basically, the input composite signal is first separated into luminance and chrominance components, and then the phase of the chrominance component signal is inverted every other frame to produce a phase-compensated signal. The phase-compensated signals Y_m and C_m are encoded through the interframe coder as follows. Shown in Fig. 2 is the preprocessor, consisting of an orthogonal transformer for T-mode and a comb filter for S-mode. In the orthogonal transformer (OTF), a pair of lines L_{2m} and L_{2m+1} are used to yield luminance Y_m and phase-compensated chrominance C_m components by the following second-order orthogonal transformation.

$$\begin{bmatrix} Y_m \\ C_m \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ (-1)^n & -(-1)^n \end{bmatrix} \cdot \begin{bmatrix} L_{2m} \\ L_{2m+1} \end{bmatrix} \qquad (1)$$

where n is a frame number and the operation $(-1)^n$ corresponds to the phase inversion to be made every other frame.

At the decoder, the inverse transformation of (1) is made to reproduce the composite signal L_{2m} and L_{2m+1}.

$$\begin{bmatrix} L_{2m} \\ L_{2m+1} \end{bmatrix} = \begin{bmatrix} 1 & (-1)^n \\ 1 & -(-1)^n \end{bmatrix} \cdot \begin{bmatrix} Y_m \\ C_m \end{bmatrix}. \qquad (2)$$

It should be noted that the phase compensation represented by (1) and (2) produces no distortion, because it is a reversible process.

In the comb filter for S-mode, chrominance (C) and luminance (Y) components are separated using a $1H$ delay circuit and a bandpass filter. The phase of the chrominance signal C

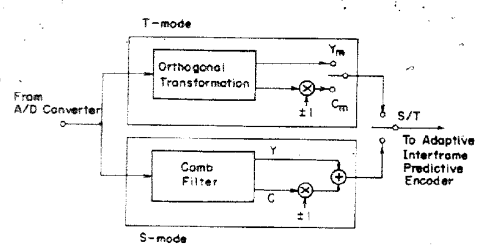

Fig. 2. Preprocessor for subcarrier phase compensation.

is inverted every other frame and added to the low-pass filtered Y to yield a phase compensated color TV signal. Strictly speaking, this process gives rise to slight degradation in color fidelity because it is irreversible.

B. Adaptive Prediction Interframe Coding

1) *Prediction Function:* A block diagram of an adaptive interframe predictive encoder is shown in Fig. 3. The coder has two predictors. One is an interframe predictor, $P_1(z)$, and the other is an intraframe predictor, $P_2(z)$.

Using the Z-transform representation

$$P_1(z) = z^{-F}, \quad \text{for } T \text{ and } S\text{-mode}$$

$$P_2(z) = \begin{cases} z^{-3}, & \text{for } T\text{-mode} \\ z^{-2}, & \text{for } S\text{-mode} \end{cases}$$

where z^{-3} means three sample delay and z^{-F} one frame delay. A more nearly optimum prediction function could be determined [1], [17], but the simple function $P_2(Z)$ above is used for the sake of hardware simplicity. The simplification results in an increase in the amount of information by only 5 percent. The adaptive prediction is made by exclusively choosing one of the two predictors.

The choice of the prediction functions is made according to the quantized prediction error amplitude. If the prediction error is smaller than a threshold level TH, the present switching signal is continued in order to hold the same prediction. When the quantized prediction error amplitude exceeds TH, the prediction switching signal is inverted in polarity by EX-OR gate,

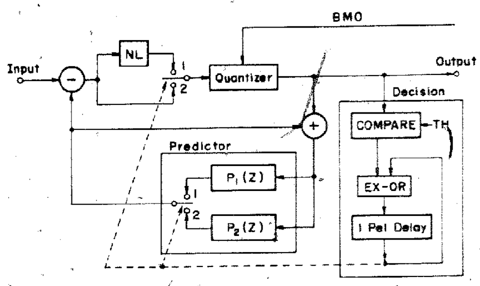

Fig. 3. Adaptive interframe predictive encoder.

and the prediction is switched to the other. Thus, the choice is made on a sample-by-sample basis. For still parts of the pictures, the interframe prediction error is almost zero. Conversely, it is larger than intraframe prediction error for the moving parts. On account of these prediction error properties, the adaptive prediction algorithm gives nearly optimum prediction, since the interframe coding is used for still parts of pictures and the intraframe coding is used for moving parts. It is not necessary to transmit the prediction switching information because it is included in the magnitude of prediction error which is transmitted. The optimum value of *TH* depends upon picture contents, but a value around 10/256 is nearly optimum for most pictures.

This adaptive algorithm provides better coding performance than the former NETEC-22H algorithm [11] based upon the third previous sample difference of the frame difference technique. Computer simulation results are shown in Fig. 4, where a picture is panned at a speed of 0–11 pels/frame. Fig. 4 shows that the coding performance of this algorithm is high compared with that of interframe, intraframe, and the third previous sample difference of the frame difference coding techniques. Particularly, the improvement by the adaptive predictions is prominent for pictures panned at high speeds. As the speed becomes higher, the probability of interframe prediction being selected tends to be 0.5. A theoretical analysis based upon a simple signal model gives a result which agrees with the simulation study [18].

2) *Nonlinear Function (NL):* A nonlinear function *NL* is applied to the frame difference signal, taking a key role in improving the coding performance. The number of significantly changed pels caused by signal noise and quantization noise is greatly decreased by applying *NL*. The nonlinear function has an input-to-output relation as shown in Fig. 5. The transfer gain is less than unity for small input amplitude and unity for large amplitude. When the transfer gain of the frame difference path is less than unity, the loop transfer function of the interframe coding has a recursive type low-pass characteristic along the temporal axis. The low-pass filtering suppresses random noise in the input signal which would otherwise cause unwanted changed picture elements even for still parts of pictures. Recurring quantization noise is also suppressed through the temporal low-pass filtering.

The nonlinear characteristic as shown in Fig. 5 is useful because small amplitude noise in the still parts of pictures is suppressed through the temporal low-pass filtering, and large amplitude frame differences caused by motions are not affected at all. The nonlinear function causes distortion in pictures with small brightness change from frame to frame. The effect is hardly seen in *NL*1 and *NL*2, although it may be perceived in *NL*3.

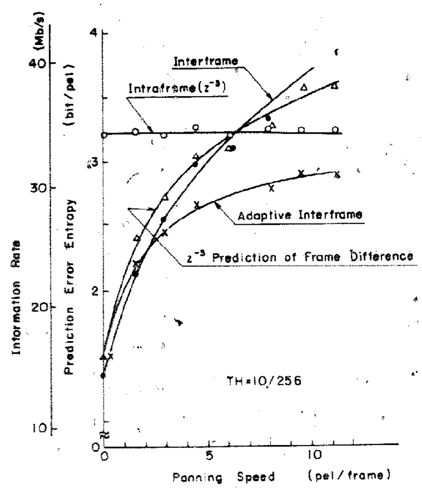

Fig. 4. Comparison of four kinds of prediction methods.

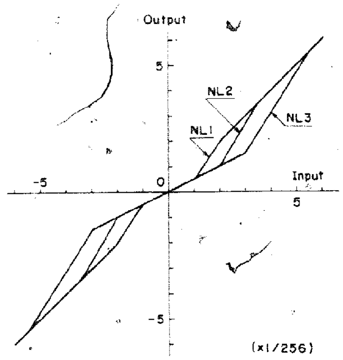

Fig. 5. Input-to-output characteristics of nonlinear function *NL*.

The influence of the nonlinear function on static performance measures such as differential gain (DG) and differential phase (DP) is small. The results obtained show that DG is 2 percent and DP 1.0°. The signal-to-noise ratio measured for a 15.734 kHz sinusoidal wave input is 55 dB weighted, which meets the broadcast picture quality requirement.

3) *Quantizing Characteristics:* Since the amount of information generated by the interframe coder varies with picture contents, it should be controlled to prevent the buffer memory from overflowing and underflowing. The information generation rate can be controlled by changing quantizing characteristics as well as other parts of the algorithm. In other words, the information rate can be controlled at the cost of picture quality. Generally, if the quantizing step size is doubled, the information rate is decreased by 1 bit/pel. In this encoder, eight quantizers, $Q0$–$Q7$, are used to provide a gradual control,

TABLE I
QUANTIZING CHARACTERISTICS OF $Q0$, $Q1$, AND $Q2$

	Quantizer Output Levels (x 1/256)
Q0	0 1.0 2.0 3.0 4.0 5.0 6.0 7.0 8.0 9.0 10.5 ..... 127.5
Q1	0 1.5 3.0 4.5 6.5 8.5 10.5 12.5 14.5 16.5 ..... 127.5
Q2	0 2.5 5.0 7.5 10.5 13.5 16.5 .... 127.5

and one of them is adaptively selected by feedback control using buffer memory occupancy (BMO) values.

The quantizing characteristics of the finest three quantizers $Q0$, $Q1$, and $Q2$ are shown in Table I. The quantizing step size of the finest quantizer $Q0$ is 1.0/256 for small amplitude, providing an 8 bit PCM equivalent quality for small amplitude change. Those of $Q1$ and $Q2$ are 1.5/256 and 2.5/256, respectively. The number of quantizing levels is 61 in $Q0$. The maximum output level is chosen to be 127.5/256, which is large enough to prevent slope overload.

SNR representation of picture quality for these quantizing characteristics is made by assuming that the SNR is given by the minimum quantization step size, although the quantizing characteristics are nonuniform. When the peak-to-peak luminance amplitude (V_{pp}) is set to be 142/256, the relation between SNR and the quantizing step size ($S/256$) can be expressed by the following equation.

$$\begin{aligned} \mathrm{SNR} &= 20 \log_{10} (V_{pp}/N_{\mathrm{rms}}) \\ &= 53.8 - 20 \log_{10} S \quad \text{(dB unweighted)} \end{aligned} \tag{3}$$

where

$$N_{\mathrm{rms}} = S/\sqrt{12}.$$

According to the equation, quantizers $Q0$, $Q1$, and $Q2$ provide the SNR values of about 54, 50, and 46 dB unweighted, respectively.

C. Variable Wordlength Coding

In the interframe predictive encoding, the prediction error is quantized and coded into a code with 6 bits/sample. The codes are transformed into reduced bit rate data through the variable wordlength coder.

The function of the variable wordlength coder is block coding for the significant pel positions and variable length coding for the significant pel amplitude. The variable length code has two sets of codes. One is a variable wordlength code set with the code length ranging from 1 to 12 bits. The unit length code is assigned to insignificant pels. This variable length code set is similar to a Huffman code and provides information amounts almost equal to entropy values. The other is a fixed-length code set with 6 bit length, which is used in conjunction with the variable one in order to avoid the continuation of long codes. Transition between the two code sets is determined by comparing the prediction error amplitude with a certain threshold level. The transition information is not needed at the receiver. The use of these two code sets is particularly effective for encoding pictures with violent motions or detailed patterns because long codes are otherwise generated in these pictures.

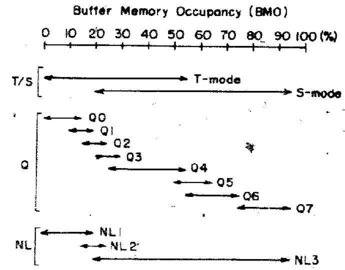

Fig. 6. Mode control diagram.

Furthermore, continuation of the insignificantly changed picture element code is deleted by block coding. Picture elements are divided into unit blocks. If all the samples in the unit block are changed insignificantly, the unit block is deleted. The positions of the deleted unit blocks are represented by block address codes. The use of block address codes is quite useful for efficiently encoding still parts of pictures.

Thus, the variable wordlength coder is capable of reducing the amount of encoded data with high efficiency, for all still, moderate, and active pictures.

D. Coding Parameters Control

Since the rate of encoded information varies with picture content, it is necessary to smooth out the irregular data generation by using the buffer memory. The capacity of the buffer memory is determined from the propagation delay time tolerated by communication links. The buffer capacity used is about 1 Mbit.

The control of the coding parameter combination among quantizers, nonlinear functions, and T/S-mode is made according to buffer memory occupancy (BMO) values. The control diagram is shown in Fig. 6. The BMO value is expressed in terms of precentage occupancy. The combination of $NL1$ and $Q0$, for instance, is used for the BMO value ranging from 0 to 15 percent. This combination provides the best picture quality, equivalent to the 8 bit PCM accuracy. As BMO increases, the coarser quantizers are used. That is, the significance determination level varies depending upon BMO values. Transition between T-mode and S-mode operations occurs at the BMO value of about 55 percent. When S-mode operation occurs for large BMO values, only a slight degradation may be perceived in the horizontal resolution, although it is rare. The combinations of the coding parameters are determined, after examination for varieties of pictures, so that the relation between picture quality and the rate of encoded information can be made optimum.

III. NETEC-22H CODING PERFORMANCE FOR BROADCAST TV PROGRAMS

A. Statistical Measurements

Coding performance is often represented by the resulting SNR. However, in interframe coding, SNR varies with picture contents to be encoded because quantizing characteristics are changed according to the rate of generated information. As there is no qualified objective evaluation method for motion

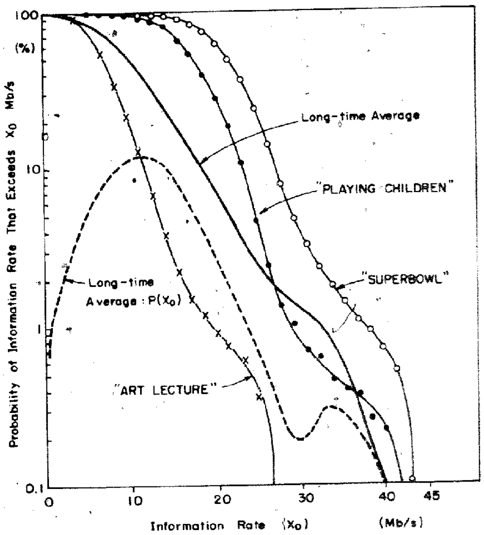

Fig. 7. Statistical information rates ($Q1$, $NL1$). $P(X_0)$ is a long-time average probability density function.

video at present, coding performance is evaluated from source signal statistics about information rate and coding parameter statistics.

In order to estimate information content, the coder parameters are fixed so that a certain picture quality is provided, then the probability distribution of the rate of generated information is measured for actual television signals for many hours. From the probability distribution thus obtained, the probability of providing the picture quality at a given transmission bit rate is calculated. In the measurements, information rates per frame were measured at the output of the variable word-length coder (point A shown in Fig. 1) with coding parameters fixed; i.e., no feedback control. One quantizing characteristic, $Q1$, which provides an SNR of 50 dB unweighted slightly better than broadcast picture quality for long-haul transmission, is used constantly in conjunction with $NL1$ under T-mode operation. Measurements were made for about 36 h applying off-the-air TV signals from four different broadcasting stations, received by a TV receiver. The SNR of the received TV signals was about 40 dB, and random noise was perceptible.

The cumulative probabilities of information rates obtained for three typical programs and the long-time average distribution are shown in Fig. 7. Here, the information per frame is multiplied by 30 (frames/s), and the horizontal axis is expressed in Mbits/s. The three fine solid lines are examples showing that generated information rates are considerably different because of different image activities.

"Art Lecture" is an example of inactive pictures, in which most of the time is occupied by still or almost still pictures. This cumulative probability distribution shows that 99 percent of the program material can be transmitted at 20 Mbits/s by using $Q1$ ($NL1$) because the probability accumulated from 0 to 20 Mbits/s is 99 percent. The average information rate in this program is as low as 8.2 Mbits/s. In "Playing Children," many children are always running around, and cameras are following them or frequently switched. Because of the motion of the children and of the camera, the cumulative probability at 20 Mbits/s decreases to 73 percent. However, the cumulative probability at 30 Mbits/s is 99 percent. "Super Bowl '79" was the most active program encountered in the measurement. This program is a football game, which was broadcast in Japan in January 1979 after reception via satellite relay and was recorded in a U-matic VTR from the off-the-air signal. For this program, the cumulative probability at 20 Mbits/s is 35 percent; however, it is still 96 percent at 30 Mbits/s. The average information rate is as high as 23.5 Mbits/s, being 15.3 Mbits/s higher than that of "Art Lecture." The above data indicate the range of broadcast TV signal source variations.

The long-time average probability density function $P(X_0)$ over 36 h is shown with a broken line in Fig. 7. The mean value of this average distribution is no more than 15.3 Mbits/s. The probability of the rates greater than the average value rapidly decreases, and a small peak is seen at around 30–40 Mbits/s. This small peak is due to scene changes. It is interesting that the shape of this distribution is quite similar to that obtained by Seyler [19], where the number of significant frame difference pels was used as a measure of information rate.

The bold solid line, representing the cumulative probability of the average distribution, shows that 93 percent of the program materials can be encoded at 20 Mbits/s and 99 percent of them can be encoded at 30 Mbits/s when a quantizing characteristic $Q1$ ($NL1$) is used. With a quantizer coarser than $Q1$, the rate of generated information becomes smaller. Therefore, when the adaptive control is operated, the residual probability of 7 percent at 20 Mbits/s and 1 percent at 30 Mbits/s will be transmitted by using coarser quantizers providing slightly lower SNR.

B. Information Generation Control

In the encoder under normal operation, quantizing characteristics are adaptively changed in order to control the rate of the information generation. As "Super Bowl '79" is an extreme case and gives a very large information rate, it is used as an example to show the performance under feedback control. The probablity distributions of information rates with three different quantizers, $Q0$ ($NL1$), $Q1$ ($NL1$), and $Q2$ ($NL2$), are compared in Fig. 8. The average values with the three quantizers are 28.6, 23.5, and 16.7 Mbits/s, respectively. Then the difference between $Q0$ and $Q1$ is about 5 Mbits/s and that between $Q1$ and $Q2$ is about 7 Mbits/s. These results show the effectiveness of changing quantizing characteristics to control information generation. It can be calculated from Fig. 8 that when the adaptive coder parameter control is made at the transmission bit rate of 30 Mbits/s, 82 percent of the program is transmitted with $Q0$, 14 percent with $Q1$, and 3 percent with $Q2$, respectively. The residual percent is transmitted by using coarser quantizing characteristics than these three. About half of the residual 1 percent is considered to be due to scene changes.

At the moment of scene changes, frame-to-frame correlation will be lost in general and, therefore, the intraframe prediction will be used in the adaptive predictive encoder. The in-

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.

WHAT WILL YOU BUILD? | sales@docketalarm.com | 1-866-77-FASTCASE