

Design of Ring Convolutional Trellis Codes for MAP Decoding of MPEG-4 Imagery

Srijidra Mahapakulchai[‡] & Robert E. Van Dyck[†]

[‡] The Pennsylvania State University, University Park, PA

[†] National Institute of Standards and Technology, Gaithersburg, MD

sxm175@psu.edu, vandyck@antd.nist.gov

Abstract—Previously, we proposed a trellis coded modulation system using CPFSK and ring convolutional codes for transmitting the bits generated by an embedded zerotree wavelet encoder. Improved performance is achieved by using maximum *a posteriori* (MAP) decoding of the zerotree symbols. In this paper, the optimal ring convolutional trellis codes are determined for MAP decoding. The CPFSK transmitter is decomposed into a memoryless modulator and a continuous phase encoder over the ring of integers modulo 4; the latter is combined with a polynomial convolutional encoder over the same ring. In the code design process, a search is made of the combined trellis, where the branch metrics are modified to include the source transition matrix. Simulation results of image transmission are provided using the optimized system.

I. INTRODUCTION

Binary convolutional codes have traditionally been designed by considering the asymptotic performance, where maximum likelihood (ML) and MAP criteria yield the same results [1]. Typically, a search is made over all noncatastrophic convolutional codes of a given constraint length in order to find the one with the minimum squared Euclidean distance, d_{free}^2 . The resulting code yields the best performance at high signal-to-noise ratios (SNRs). For bandwidth constrained channels, one can achieve a good coding gain by incorporating the convolutional code into a trellis-coded modulation (TCM) [2] system, usually using a mapper to convert bits to *m*-ary symbols. Rimoldi [3] showed how to decompose a continuous phase modulation system into a continuous phase encoder (CPE) and a memoryless modulator, thereby allowing simplification of the code design process.

A. Joint Source-Channel Coding

For many personal communication systems, low power in battery-operated devices and a fading environment caused by user mobility makes the high SNR assumption of dubious validity. These systems also have limited computational power, precluding long constraint length codes and sophisticated iterative decoding schemes. In such

cases, Shannon's separation theorem is less applicable, and gains can be achieved by using joint source-channel coding. One powerful method is to utilize the residual redundancy in the source [4] [5] [6] [7] [8], by designing a source-controlled channel decoder. For a discrete channel, we employed [9] the hidden Markov model (HMM) viewpoint of Miller and Park [5] to develop a MAP decoder for the MPEG-4 codec [10].

Kroll and Phamdo [6] [7] studied the transmission of data from a binary Markov source over the continuous AWGN channel. After first using Ungerboeck codes, they designed trellis codes especially for this source and MAP decoding. Al-Semari *et al.* [8] examined sequence MAP decoding for Gaussian and Rayleigh channels. Recently [11], we used quaternary continuous phase frequency-shift keying (CPFSK) TCM in conjunction with ring convolutional codes for MPEG-4 image transmission over the AWGN and Rayleigh channels. Since the image coder outputs quaternary symbols, the source, the channel encoder, and the continuous phase encoder are all in the same algebraic framework. The trellis codes were based on those designed by Yang and Taylor [12], although they were converted from recursive systematic form to polynomial form to allow MAP decoding.

Here, we design ring convolutional codes explicitly for the CPFSK TCM system with MAP decoding. For a given constraint length, we find the codes that provide the best performance for symbol error rates in the range of 10^{-2} to 10^{-4} . Since the high SNR assumption is no longer valid, a methodology is proposed that yields optimal codes. In the process, the source transition matrix is incorporated into the branch metrics used in the trellis search.

B. MPEG-4 Image Transmission

The error resilient transmission of MPEG-4 still imagery is the ultimate goal of this work. A block diagram of the entire system, including the modulation and ring convolutional channel coding, is given in Figure 1 of [11]. Here we provide a very brief description of the image compression and decompression algorithms. Further discus-

sion, including the design of the packetization scheme, can also be found in [9].

The MPEG-4 embedded zerotree wavelet (EZW) algorithm [10] first performs a discrete wavelet transformation of the image. The lowest frequency subband (LFS) is scalar quantized and then DPCM predicted. The higher frequency subbands (HFS) are also scalar quantized and then zerotree scanned. There are two modes of processing of the HFS: 1.) single quantization and 2.) multi-quantization [13, pp. 130-133]. Here we concentrate only on the multi-quantization mode, which employs bit plane encoding and multiple zerotree scans. The bits generated by compressing the lowest frequency subband are placed into a fixed-length packet, while the bits from the higher frequency subbands are placed into a number of variable-length packets. Overhead information is included so that each packet can be independently processed. The packetized bit stream is converted to quaternary symbols and input to the channel encoder and modulator. The convolutional encoder is reset for each packet. Note that only the zerotree symbols in each HFS packet can be MAP decoded; the remaining data and overhead must be ML decoded.

II. TRELLIS CODED CPFSK

A. Ring Convolutional Codes

Rimoldi [3] showed that one can decompose CPFSK with modulation index $h = \frac{1}{M}$ into the continuous phase encoder and a memoryless modulator (MM), where the CPE is in the form of a convolutional encoder over the ring of integers modulo-M, Z_M . Thus, an external convolutional encoder (CE) over Z_M is a natural way to combine channel coding and the CPE, without using a mapper. Yang and Taylor [12] proposed a TCM system based on the combination of an external recursive systematic ring convolution encoder and the CPE. Their results show that this coding scheme provides better performance than other comparable systems. Rimoldi and Li [14] also examined this combination, including the design of systematic codes for partial response continuous phase modulation.

In this paper, polynomial (nonsystematic) ring convolutional codes over Z_M are applied for two reasons. Firstly, we observe that the best noncatastrophic polynomial ring CEs provide larger d_{free}^2 than the best systematic ring convolutional codes found in [12]. Secondly, the trellis diagram of the overall encoder (CE+CPE) allows one to easily incorporate the source transition matrix (STM) into the branch metrics, in order to perform MAP decoding [15] at the receiver. The details of this procedure are given in [11]. Since the polynomial CE over Z_M can be catastrophic, a

test is required to determine if the resulting polynomial encoder over the ring is noncatastrophic. This task can be achieved by using the following theorem.

Theorem: “A polynomial encoder $G(D)$ over the ring Z_M , where $M = p^m$ and p is a prime, is catastrophic if and only if, when the coefficients of the polynomials in $G(D)$ are each reduced modulo p , the resulting polynomial encoder over the finite field $GF(p)$ is catastrophic.” [16]. For instance, a polynomial encoder $G(D) = [D + 1 \ 3D + 1]$ over the ring Z_4 is catastrophic, since all coefficients of the elements of $G(D)$, reduced modulo 2, result in the polynomial encoder $[D + 1 \ D + 1]$ over the finite field $GF(2)$. From [17], it is easy to see that the encoder $[D + 1 \ D + 1]$ is catastrophic; therefore, so is $G(D)$.

B. Code Search Results

In a Gaussian channel, the minimum squared Euclidean distance (SED), also known as d_{free}^2 , plays an important role in code design, and it can be computed by using the Viterbi algorithm [15]. The observation interval length is chosen to be large enough to ensure that d_{free}^2 is found for each code. It is defined by

$$d_{free}^2 = \min_{\mathbf{u} \neq \mathbf{u}'} d^2(\mathbf{u}, \mathbf{u}'). \quad (1)$$

Without loss of generality, the distance is computed between the all zero sequence, \mathbf{u} , and the set of all other sequences. The metric $d^2(\mathbf{u}, \mathbf{u}')$ is the squared Euclidean distance between two sequences, and it can be simplified to

$$d^2(\mathbf{u}, \mathbf{u}') = \sum_{n=0}^N d_n^2(\mathbf{u}, \mathbf{u}'), \quad (2)$$

where $d_n^2(\mathbf{u}, \mathbf{u}')$ is the incremental squared Euclidean distance (ISED) and N is the number of observation intervals. The ISED can be calculated by using the input of the MM. The details of this derivation can be found in [12].

For a given constraint length, the codes having maximum d_{free}^2 provide the best performance at low symbol error rates. This fact is based on a union bound analysis. When the SNR is large, the bound is dominated by one term consisting of d_{free}^2 . Even though the ultimate goal is to find the best code for a particular STM at high symbol error rates, for comparison purposes, the code search is first done at a low symbol error rate; this is the standard ML criterion. Table I shows the best rate $\frac{1}{2}$ polynomial ring convolutional encoders and their maximum d_{free}^2 for the given total number of states of the overall encoders.

Note that these nonsystematic ring convolutional encoders give larger d_{free}^2 than the systematic ring convolutional encoders found in [12]. In order to determine the

S_t	G	d_{free}^2 (YT)	d_{free}^2 (ML)
4	$[3D+3, 3]$	3.15	3.58
8	$[2D^2+3D+3, 3]$	4.09	4.36
16	$[3D^2+D+2, 2D+3]$	5.15	5.79
32	$[2D^3+D^2+D+3, D+2]$	6.00	6.03
64	$[D^3+D^2+2D+1, D^2+2D+2]$	6.42	7.45
128	$[D^3+D^2+3, 2D^3+D^2+2D+2]$	7.60*	8.00

TABLE I

THE BEST RATE $\frac{1}{2}$ NONSYSTEMATIC RING CONVOLUTIONAL ENCODERS OVER Z_4 FOR 4-ARY CPFSK WITH $h = \frac{1}{4}$. S_t DENOTES THE TOTAL NUMBER OF STATES IN THE OVERALL ENCODER. YT IS THE BEST SYSTEMATIC CODE FROM [12], WHILE ML IS OUR BEST CODE DESIGNED FOR MAXIMUM LIKELIHOOD DECODING. * INDICATES THE SEARCH WAS INCOMPLETE.

best codes for MAP decoding, the squared Euclidean distance in Eq. (2) is modified in such a way that it includes the source transition matrix

$$d_{map}^2(\mathbf{u}, \mathbf{u}') = \sum_{n=0}^N (d_n^2(\mathbf{u}, \mathbf{u}') - N_0 \ln \frac{P(u_n|u_{n-1})}{P(u'_n|u'_{n-1})}), \quad (3)$$

where the sequence $\mathbf{u} = (u_0, u_1, \dots, u_N)$. When the conditional probability $P(u_n|u_{n-1}) = P(u'_n|u'_{n-1})$, Eq. (3) reduces to Eq. (2). To determine the minimum of d_{map}^2 for each code, we again use the Viterbi algorithm. Note that the noise power, N_0 , in Eq. (3) must be specified. This fact implies that the optimized code is the best code at a particular SNR. In this work, we use this new distance to find the best codes optimized for the source transition matrix from the “Lena” image, which is shown below.

0.2391	0.1674	0.4090	0.1845
0.2510	0.1241	0.4157	0.2092
0.1262	0.0908	0.5407	0.2423
0.1506	0.1136	0.5470	0.1888

Table II gives the best codes designed for MAP decoding at a symbol error rate (SER) of approximately 10^{-3} . This performance corresponds to an E_s/N_o ¹ of 4.5 dB for 4 and 8 states and 3.5 dB for 16 and 32 states. With 4, 8, and 16 states, the best codes are the same for MAP and ML decoding. For 32 states, a different code is found; comparing this code to the best ML code shown in Table I and using MAP decoding for both, a slight improvement of about 0.1 dB is obtained at low SNR. Note that d_{free}^2 of this new code is less than that of the ML code, and so

¹Energy per symbol to noise power spectral density.

S_t	G	d_{free}^2 (YTE)	d_{free}^2 (MAP)
4	$[3D+3, 3]$	3.15	3.58
8	$[2D^2+3D+3, 3]$	4.09	4.36
16	$[3D^2+D+2, 2D+3]$	5.15	5.79
32	$[D^2+3D+1, 2D^2+3D+2]$	5.32	5.67

TABLE II

THE BEST RATE $\frac{1}{2}$ NONSYSTEMATIC RING CONVOLUTIONAL ENCODERS OVER Z_4 FOR 4-ARY CPFSK WITH $h = \frac{1}{4}$. YTE IS YANG AND TAYLOR’S EQUIVALENT CODE, WHILE MAP IS OUR BEST CODE DESIGNED FOR MAP DECODING.

at high SNR, the performance is slightly (negligibly) less. Comparing this MAP-designed code to Yang and Taylor’s equivalent code², one sees a gain of about 0.2 to 0.3 dB at symbol error rates ranging from 10^{-2} to 10^{-3} , again both using MAP decoding.

III. SYMBOL ERROR RATE RESULTS

The source transition matrix of the “Lena” image, shown in the previous section, is used here. Figure 1 (top) shows the symbol error rate vs. E_s/N_o for 4-CPFSK with two different ring convolutional encoders using ML and MAP decoding. For comparison purposes, MSK without any channel coding is also included. The sixteen and thirty-two state codes provide a performance improvement of about 0.3 to 0.5 dB for MAP decoding at symbol error rates in the range of 10^{-2} to 10^{-3} . For the cases of four and eight states (not shown), the simulation results show a performance improvement of approximately 0.25 to 0.4 dB at this same range of interest. The results in Figure 1 (bottom) illustrate the performance of both the MAP and ML decoders over a Rayleigh flat fading channel. Similar improvement from MAP decoding is seen.

IV. IMAGE TRANSMISSION RESULTS

As previously mentioned, the improvement of the image quality after reconstruction is our ultimate goal. In this section, PSNR simulation results are obtained. After transmission over the channel, the received zero-tree symbols in the HFS packets with a nominal length of 500 bits are decoded using MAP decoding³, implemented by the Viterbi algorithm. The received LFS packet and the rest of each HFS packet are ML decoded. Presently, it is assumed

²The equivalent code is found by converting the given systematic code to polynomial form. This change is necessary for incorporating the source transition matrix.

³In section IV-B, we compare against ML decoding. In this case, all of the packets are decoded using the ML criterion.

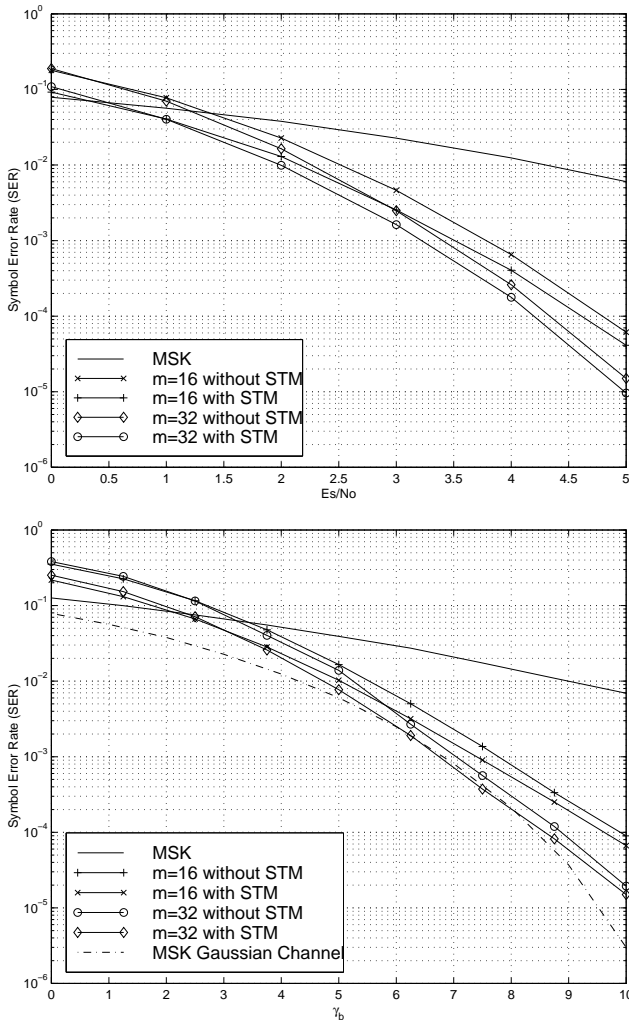


Fig. 1. (top) Symbol error rate vs. E_s/N_0 for AWGN channel. (bottom) Symbol error rate vs. $\bar{\gamma}_b = \frac{1+\gamma}{\gamma} \frac{E_s}{N_0}$ for the fading channel. E_s/N_0 is the energy per channel symbol, while the SER is for the information symbols. Lena source transition matrix. m is the number of states in the trellis.

that the receiver knows the number of zerotrees in each packet. The 32 state code is used for both the proposed MAP system and the baseline ML system.

The PSNR results shown in Tables III and IV are for decoding strategy 1. In this strategy, the MPEG-4 decoder decodes HFS packets if they pass an outer cyclic redundancy check. Note that this check requires only 4 bits/packet so it is fairly weak. Consequently, undetected errors can be noticeable at high symbol error rates.

A. Code Choice for MAP Decoding

In this section, the performance comparison between our best code and the YTE code is obtained. Table III gives the performance improvement in terms of PSNR of the decoded image. In the range of symbol error rates from 10^{-2} to 10^{-3} (E_s/N_0 2-3 dB), the improvement is from 1.6 to 1.8 dB. One reason for this difference is that by constrain-

$\frac{E_s}{N_0}$ (dB)	PSNR (dB) MAP (YTE)	PSNR (dB) MAP (BC)	Δ PSNR (dB)
1.0	8.26	8.49	0.23
2.0	11.83	13.42	1.59
3.0	18.56	20.35	1.79
3.5	22.14	23.18	1.04
4.0	24.50	25.10	0.60
4.5	26.68	26.73	0.05
5.0	27.82	27.82	0.00

TABLE III

THE COMPARISON OF RECONSTRUCTED IMAGE AVERAGE PSNR BETWEEN OUR BEST CODE (BC) AND YANG AND TAYLOR'S EQUIVALENT CODE (YTE) USING MAP DECODING. 200 IMAGE TRIALS PER DATA POINT. STRATEGY ONE IS USED, AND IT DECODES ALL HFS PACKETS THAT PASS A PARITY CHECK. AWGN CHANNEL.

ing the code search to systematic codes, the best noncatastrophic ring convolutional codes are not necessarily found. This result is consistent with Table I in [1], which is for rate one-half binary codes. Also, converting from systematic form to polynomial form can decrease d_{free}^2 , as seen by comparing Tables I and II.

B. MAP vs. ML Decoding

Now, we determine the gains possible by using the optimal codes with MAP decoding instead of ML decoding. Table IV shows that for both the AWGN and Rayleigh channels, the use of MAP decoding and ring convolutional codes improves the image PSNR. For the Gaussian channel, an improvement of almost 0.5 to 0.6 dB can be evidently seen at SNRs from about 3 to 3.5 dB. For the fading channel, this gain is most evident at $\bar{\gamma}_b$ values in the 5 to 9 dB range, corresponding to symbol error rates in the 10^{-2} to 10^{-4} region. The improvement ranges from about 0.1 to 0.4 dB. Figure 2 shows the reconstructed image obtained using MAP decoding after transmission over the Rayleigh channel. The PSNR is 28.93 dB for MAP decoding, while the PSNR for ML decoding is 25.63 dB. This experiment indicates that while the average improvement over many trials is only 0.178 dB, as shown in Table IV, there are individual transmissions with much higher gains.

V. CONCLUSIONS

In this paper, we designed polynomial ring convolutional codes explicitly for MAP decoding of MPEG-4 imagery. Two types of comparisons were made. Firstly, the performance of these codes were compared to that of polynomial codes obtained by converting the best systematic ring convolutional codes of [12]. In these experiments, MAP decoding was used for both, and gains in PSNR of

$\frac{E_s}{N_0}$ (dB)	PSNR (dB) ML	PSNR (dB) MAP	Δ PSNR (dB)
1.0	8.487	8.490	0.003
2.0	13.264	13.418	0.154
3.0	19.766	20.353	0.587
3.5	22.700	23.183	0.483
4.0	24.812	25.103	0.291
4.5	26.295	26.731	0.436
5.0	27.626	27.822	0.196
5.5	28.595	28.595	0.000

$\bar{\gamma}_b$ (dB)	PSNR (dB) ML	PSNR (dB) MAP	Δ PSNR (dB)
1.25	6.553	6.553	0.000
2.50	7.460	7.461	0.001
3.75	9.476	9.489	0.013
5.00	13.313	13.462	0.149
6.25	18.627	19.026	0.399
7.50	23.245	23.615	0.370
8.75	26.120	26.298	0.178
10.00	27.610	27.639	0.029
11.25	28.394	28.396	0.002

TABLE IV

RECONSTRUCTED IMAGE AVERAGE PSNR USING ML AND MAP DECODING. THE TOP IS THE AWGN CHANNEL AND THE BOTTOM IS THE FLAT FADING CHANNEL. 200 IMAGE TRIALS PER DATA POINT. DECODING STRATEGY ONE.

up to 1.79 dB were found for an AWGN channel. It must be noted that the systematic codes were designed for ML decoding, and that the conversion to polynomial form is partly responsible for their reduced performance.

Secondly, the MAP-designed codes were compared to the best polynomial codes designed for maximum likelihood decoding. The former were MAP decoded while the later where ML decoded. The PSNR gains were more modest. Still, the average improvement can be 0.4 to 0.6 dB for both AWGN and Raleigh channels, with individual image trials yielding substantially better results. Additionally, for 4, 8, and 16 states, the same code is found to be optimal whether the metric of Eq. (2) or Eq. (3) is used.

REFERENCES

- [1] A. J. Viterbi, "Convolutional codes and their performance in communication systems," *IEEE Trans. on Comm. Tech.*, vol. COM-19, pp. 751-772, Oct. 1971.
- [2] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. on Info. Theory*, vol. IT-28, pp. 55-67, Jan. 1982.
- [3] B. E. Rimoldi, "A decomposition approach to CPM," *IEEE Trans. Information Theory*, vol. 34, no. 2, pp. 260-270, Mar. 1988.
- [4] M. J. Ruf and J. Hagenauer, "Source-controlled channel decoding in image transmission," *Proc. Workshop on Wireless Image/Video Comm.*, Loughborough U.K., pp. 14-20, Sept. 1996
- [5] D. J. Miller and M. S. Park, "A sequence-based approximate MMSE decoder for source coding over noisy channels using discrete hidden Markov models," *IEEE Trans. Comm.*, pp. 222-231, Feb. 1998.



Fig. 2. Reconstructed image using MAP decoding. PSNR = 28.93 dB. 32 state code; fading channel with $\bar{\gamma}_b = 8.75$ dB. The average symbol error rate is about 10^{-4} .

- [6] J. M. Kroll and N. C. Phamdo, "Analysis and design of trellis codes optimized for a binary symmetric Markov source with MAP detection," *IEEE Trans. on Info. Theory*, vol. 44, pp. 2977-2987, Nov. 1998.
- [7] J. M. Kroll and N. C. Phamdo, "Source-channel optimized trellis codes for bitonal image transmission over AWGN channels," *IEEE Trans. on Image Proc.*, vol. 8, pp. 899-912, July, 1999.
- [8] S. A. Al-Semari, F. Alajaji, and T. Fuja, "Sequence MAP decoding of trellis codes for Gaussian and Rayleigh Channels," *IEEE Trans. on Vehicular Tech.*, vol. 48, pp. 1130-1140, July 1999.
- [9] R. E. Van Dyck, "MPEG-4 image transmission using MAP source-controlled channel decoding," *IEEE JSAC*, vol. 18, no. 6, pp. 1087-1098, June 2000.
- [10] ISO/IEC JTC1/SC29/WG11, "Information technology, coding of audio-visual objects: Visual, ISO/IEC 14496-2, Committee Draft," N2202, Mar. 1998.
- [11] S. Mahapakulchai and R. E. Van Dyck, "Source-controlled channel decoding of MPEG-4 imagery using CPFSK and ring convolutional codes," *Proc. IEEE Electro-Information Technology Conf.*, Chicago, IL, June 2000.
- [12] R. H. Yang, and D. P. Taylor, "Trellis-coded continuous-phase frequency-shift keying with ring convolutional codes," *IEEE Trans. Information Theory*, vol. 40, pp. 1057-1067, July 1994.
- [13] ISO/IEC JTC1/SC29/WG11, "MPEG-4 video verification model version 11.0," N2172, Mar. 1998.
- [14] B. E. Rimoldi and Q. Li, "Coded continuous phase modulation using ring convolutional codes," *IEEE Trans. on Comm.*, vol. 43, pp. 2714-2720, Nov. 1995.
- [15] G. D. Forney, Jr, "The Viterbi algorithm," *Proc. of the IEEE*, vol. 61, pp. 268-278, Mar. 1973.
- [16] J. L. Massey and T. Mittelholzer, "Convolutional codes over rings," in *Proc. 4th Joint Swedish-Soviet Int. Workshop Information Theory* Gotland, Sweden, pp. 14-18, Aug. 27-Sept. 1, 1989.
- [17] G. D. Forney, Jr, "Convolutional codes I: Algebraic structure," *IEEE Trans. Info. Theory*, vol. IT-16, pp. 720-738, Nov. 1970.