Modified Viterbi Algorithm for Decoding of Block Codes

Zolotarev V.V.
Department 71
Space Research Institute of Russian Academy of Sciences
Moscow, Russian Federation
zolotasd@yandex.ru

Grinchenko N.N., Ovechkin G.V., Ovechkin P.V.
Faculty of Computer Engineering
Ryazan State Radio Engineering University
Ryazan, Russian Federation
ovechkin.g.v@rsreu.ru

Abstract—**The paper has considered optimal methods for decoding error correction codes which allow increasing reliability of the data transmission across noisy communication channels. The purpose of work is to solve a task to increase efficiency of decoding short block codes. The paper has suggested a very simple modification of the algorithm for decoding quasicyclic block codes built on the basis of short convolutional codes. Besides, decoding is executed by the Viterbi algorithm not taking into account a block structure of the code but decoding an ordinary infinite sequence. The analysis of the suggested decoder characteristics has been performed and complexity of its implementation has been estimated. The paper has shown that by means of the suggested Viterbi decoder for the block codes it is possible to ensure a bit error rate being analogous to the bit error rate for the corresponding convolutional code under the decoding complexity increasing only in 2..3 times.**

Keywords—data transmission system, error correction coding, Viterbi algorithm, convolutional code, block code, quasicyclic code, cyclic trellis, decoder complexity.

I. Introduction

Fundamental research challenge to increase reliability of the digital data transmission across noisy channels is solved by methods of error correction coding. Optimal decoders always finding a codeword being the closest to the received sequence are the best ones by decoding efficiency. Optimal decoders under decoding received sequence execute the full search of all possible variants of codewords, so their complexity exponentially depends on a length of the used code. At present the decoding algorithm for convolutional codes suggested by A.Viterbi in 1967 is mostly widely used among optimal algorithms for decoding [1]. The present algorithm allows executing the full search of all possible codewords and selecting that one locating at a minimum distance from the received sequence.

Under developing of communication equipment with coding systems on the basis of decoders realizing the Viterbi algorithm (VA) or other error correction methods it has very often to realize re-synchronization procedures. They consist in the fact that under completing of the transmission of some information sequence that should be sent to the recipient across the channel, zero information sequence of length equal to the encoder memory size is introduced into the encoder. This procedure called trellis termination [2] is required so that noise immunity of last information bits does not become worse because of lack of necessary code bits if code sequence transmission is terminated at the moment of the last information bit enter into the encoder. Besides, it is inconvenient that additional zero bits at the tail of information sequence of the finite length change code rate of the convolutional code. For example, under the information sequence length being equal to 100 bits, for the code with rate one half $2(K-1)$ bits should be transmitted additionally, where K is constraint code length. So, for example, in a particular case for $K=7$, 12 code bits additionally are transmitted across the channel. That means that real code rate is equal not to $R=1/2$ that will be very convenient for formation of the control actions for transmission and receiving equipment using Viterbi decoders but $R=100/212\approx0,47$ that unreasonably complicates development of communication systems.

That is why **there is a task** to develop such method for block encoding and following decoding of sequence when decoder operation is ensured under the same code rate R with the convolutional code usually expressed by a ratio of small integer numbers, i.e. block codes with convenient decoding are required. These block codes should have the same convenient simple values of code rates with the convolutional codes, for example, $R=1/2$, $R=1/3$, $R=3/4$ etc. That would significantly simplify error correction coding application with usage of VA for block codes.

The rest part of the paper is arranged as follows. In Section II we have considered short block codes based on convolutional codes and suggested modification of the Viterbi algorithm for their decoding having less computational complexity in comparison with known ones. In section III we present simulation results for submitted method over the channel with Additive White Gaussian Noise. The complexity of offered decoder for block codes is analyzed in section IV. Section V concludes this paper.

II. Method for Encoding and Decoding

Fig. 1 shows a device for encoding information sequence with suggested method transformed convolutional code with generator polynomial $g(x) = 1 + x + x^4 + x^6$ into block one with the same code rate. This method allows applying various methods for decoding, including VA, at the receiving end of the communication line [3].

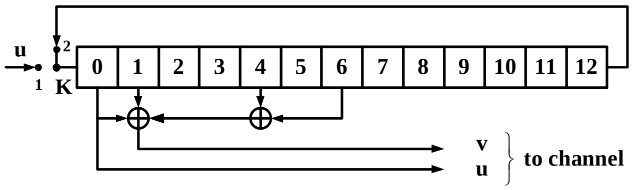

Figure 1. Encoder for the quasicyclic block code

This block code is characterized by parameters: code length n=26, information length k=13, code rate R=1/2, code distance d=5. It is based on a convolutional code with K=7, d=5 and the same generator polynomial. It is important that for this purpose the right part of the encoder register not containing taps to modulo 2 adders should be of the same or higher length than the maximum degree of the generator polynomial. It should be emphasized that such transformation of the convolutional code into the block quasicyclic code is well known and widely used for the coding technique [2, 4, 5, 6]. Additional information on these procedures often called as cyclic truncation of convolutional codes and analyzable by means of cyclic trellis can be found in [7]. References to other authors interested in the same subject are also presented there.

Above suggested device for encoding information sequence with the block quasicyclic code operates in the following way. Firstly, information symbols intended for encoding and following transmission across the channel are entered into the encoder in some way, for example, in parallel. Then the main procedure for encoding information sequence is executed. After each shift, several code bits (in the concerned example, two code bits) are encoder output. This process for the block code generation terminates when it appears at the initial state after a number of cyclic shifts of the encoder register. Then it is possible to start encoding of the next block. Note that the same scheme can be used for encoding with the non-systematic block code also easily decoding with the VA.

Let's happen at the receiving end of the communication line so that a Viterbi decoder immediately starts operating the same way as under the usual decoding the initial long (maybe, infinite) convolutional code from a random place, for example, directly from the first symbols of the quasicyclic code entering into the Viterbi decoder. After receipt and processing of the last code sub-block of the received sequence, the first sub-block, then the second etc. are again cyclically enter into the Viterbi decoder. Depending on a code length and noise level it is again possible to enter cyclically all symbols of the received sequence 2..5 and more times into the VA decoder. But since a quasicyclic code has no "beginning" then VA as well as in the convolutional code, will come to the correct (i.e. having a small number of errors) decision only after entering of the first approximately $D \approx (3..20)K$ code sub-blocks into the decoder. But then after receipt of these code sub-blocks, decisions of VA obligatory repeat also obviously with a period being equal to information sequence length. That means that a receiver should be provided from the decoder only with a part of VA decision sequence which, for example, in the case of the encoder shown in Fig. 1 has the length 13 bits. But it should be taken from the Viterbi decoder outside the first D code sub-blocks entered into the decoder where, as it as mentioned before, these decisions are generally wrong. From the other side for VA it is also well known (see [3]) that other condition of enough decision reliability consists in the fact that such correct decoder decisions are usually formed not earlier that after receipt of $(5..25)K$ code sub-blocks which absolute number also depends on a code and noise level in the channel. So, VA (or other decoder) should "accept" cyclically several similar (!) received sequences and then transmit decoded information bits locating approximately in the middle of this cyclical sequence of the VA decoder decisions to the receiver because decisions locating close to the place of the next code sub-block receipt and close to the place of the first code sub-blocks receipt are unreliable in comparison with potential possibilities of the used code.

It is obvious that correcting ability of the block code keeps almost the same as well as the convolutional code has if the block code length in comparison with the convolutional code constraint length appears greater in 10..30 or more times. Under shortening of the block code and under keeping the generator polynomial characteristics of the block code, of course, should become worse. However, simplicity and habitualness (of course, conventional) of decoding with usage of VA under keeping of the code rate for the basic convolutional code allows considering that optimal block codes decoding based on the VA allows practically solving all tasks for errors correction more easily.

III. Characteristics of Decoding Block Codes on the Basis of the Viterbi Algorithm

Fig. 2 shows simulation results for above mentioned methods for encoding and decoding block quasicyclic codes with code rate R=1/2 on the basis of classic VA. Along the vertical axis bit error rates (BER) are indicated for described block codes and also word error rates (WER) for the block codes are shown. Along the horizontal axis signal to noise ratio per bit E_b/N_0 is indicated in the channel with Additive White Gaussian Noise.

Curve VA7 corresponds with the Viterbi decoder for the standard convolutional code with K=7, but analogous curve VA15 – for convolutional code with K=15. Such decoder has already been easily realized on the contemporary integrated circuit. It is important to remember that firstly so long code was used under small code rate as far back as last millennium for the NASA project "Cassini". For other curves designations XK-n are accepted. X is corresponding with a type of error rates which curve represent: B – BER for the block modification of convolutional codes, W – WER for the same codes; K is the constraint length of the code chosen as a base for the block one (7 or 15); n is the length of the block quasicyclic code (40, 100 or 200 bits).

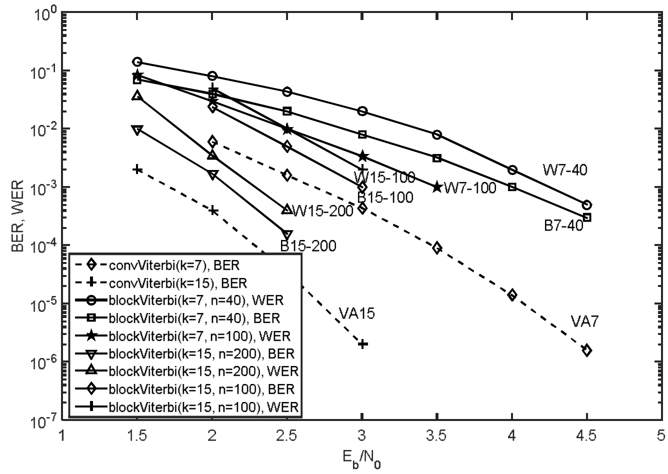

Figure 2. The performance of decoders for convolutional and block codes over the channel with Additive White Gaussian Noise

Lets discuss characteristics of the suggested method. The BER and WER curves corresponds with the expected level of error probabilities. As it was mentioned before, it is necessary that lengths of constructed quasicyclic block codes significantly exceed lengths of basic convolutional codes, besides, with increase of the parameter *K* this difference should raise in order that BER of VA decoder for block codes are close to BER of VA decoders for basic convolutional codes. For example, curve VA7 almost coincides with BER of VA for the block code of length *n*=100 with same generating polynomial. So, this curve should not be drawn separately. And under using of the same convolutional code in the block code with *n*=40 difference between parameters *K* and *n* has been small. So, characteristics of decoder for this block code with the same value *K*=7 as in the previous case become less: both curves B7-40 and W7-40 are located significantly higher than curve W7-100 of decoder for the first code.

Analogous ratio of characteristics is also seen in the case of the block codes based on convolutional code with *K*=15. For the block codes BER for their decoders will be insignificantly different from the performance of VA decoder for the convolutional code only under length *n* exceeding evidently 300..400 bits. BER and WER curves of decoders for the block codes are also shown in Fig. 2 for *n*=100 and *n*=200. They show that under decreasing of *n* performance of the block code with decoding according to VA noticeably decrease as it should be for any block code comparing with the convolutional one under their comparable lengths.

Let's note that block codes with the understandable decoding with the Viterbi algorithm similarly to its convolutional prototype are easily decoded under transfer to the VA variant with reduced state research. However, it is necessary to set its parameters very exactly.

It is also appropriate to remind that all block modifications of convolutional codes are easily included into various parallel and sequential concatenated codes. It stops a lot of issues of development and usage of various block codes which are often declared as revolution in the technique of coding and decoding without any reasons (for now!). Analysis of some methods of such "level" is shown in [8]. Comparison of represented polar codes [9] and other codes has a preliminary nature because of very restricted actual materials regarding these codes declared as perspective. It is shown there that VA and multithreshold decoders (MTD) even without any adaptations to terms of comparison are close to polar codes by efficiency that in general becomes not a very complicated task. All declared advantages of polar codes [9] under small lengths of blocks and, as a consequence, theoretically their weak performance are easily achieved by algorithms of the type VA, MTD and concatenated schemes with their participation under very low complexity [3, 5, 8, 10].

IV. COMPLEXITY OF BLOCK MODIFICATION FOR THE VITERBI ALGORITHM

Above mentioned principle of the block VA construction has shown that complexity of represented decoders of the brute-force type, and it should not be forgotten, becomes rather acceptable for a lot of applications because convolutional VA versions were comprehensively studied some decades ago, and block VA modification is not very different from the basic convolutional one.

Also it is useful to mention some peculiarities of the block VA modification and its possibilities. Necessity for each code block to decode a repeated sequence of some longer length than the block size does not complicate the algorithm very much. In the most cases only decoding of the "convolutional" code can be enough which length, for example, in the above represented experiments never exceeds a size of the block code more than in 5 times. Usually within the experiment it was possible to be restricted by threefold difference in code lengths and this value can be more reduced approximately in 1,5 times.

Under software implementation significant computational efforts are required for support of survivor paths which number, as it is known, increases exponentially with increase of the convolutional code constraint length *K*. Besides, saving of computations for reformation of these paths at each decoding step is possible due to double growth of the memory used for storing of references to previous positions of survivor paths. Also saving of paths memory is possible that gives a possibility to realize block VA for *K*~28 or even greater values. But computational efforts increase in 2..4 times. Finally, let's pay attention to the fact that block VA should only once trace back the best path after termination of the whole decoding procedure and choose an average part from the sequence of its decisions as it was indicated in the algorithm description in details. It also reduces the computational complexity.

Most of mentioned circumstances was considered and analyzed within the process of simulation for block and convolutional versions of VA both for the similar and various parameters for constructed codes. But deep procedures for computations optimization in order to increase decoding rate have not been executed. It is obvious that this useful task deserves additional consideration and attention. It appeared that for typical conditions performance of decoders for convolutional and block codes has not been very different. For the block version of VA decoder for code with K=7 with using of personal computer with Core-i7 processor on clock frequency 3 GHz decoding in the convolutional classical variant occurred with throughput about 100 Kbps, and in the block version – 60 Kbps. Under using block codes based on the convolutional code with K=15 decoder throughput was equal to 1,0 Kbps and in its ordinary convolutional variant – about 600 bps. These results testify to a good level of throughput for software block VA version. Main possibilities to improve the developed version of the block VA have been considered above.

Finally, let's indicate that the similar decoder throughput for convolutional versions of VA and their new block modifications under simulation testify that complexity of the block VA remained being equal to the order 2^K additive equivalent operations per information bit. This follows from description of the suggested algorithm which actually coincides with the base method. Because of this, it is reasonable to pay attention that methods of "cyclic truncation of tails" under transfer from convolutional to block codes, including methods described in [7], are characterized by the fact that decoders for suggested block modification of codes have complexity very close to 2^{2K}. It is obvious that so impressive complexity of the modification for the ideologically (not algorithmically) simple method is absolutely unacceptable. Authors of such methods and their evaluations rightly note that significant simplifications are possible. Some of such methods for simplification of very complicated decoder for the block codes have been suggested by them but there are no evaluations for complexity of such simplifications in the convenient form for interpretation.

V. CONCLUSION

The suggested in this paper block variant of VA is actually enough simple and very clear modification of the classical VA for convolutional codes and has complexity as it is shown in the description evidently exceeding complexity of the base algorithm not more than in 2..3 times. It is possible that some its modifications will become simpler. It determines all its wide possibilities for usage in multipurpose communication systems.

Note it is likely that at present the highest characteristics in coding gain and complexity can be provided now only by decoders with direct control of metrics. Various MTD algorithms, all VA modifications and also divergent decoding schemes [11] developed on their basis belong to them. All these methods are combined by simplicity and uniformity of computations and also property of exact measurement for distances of these decoder decisions up to received sequence which other widespread decoders do not have. We suppose that if a decoder does not take into consideration and does not measure an exact distance up to the received sequence, that such error correction algorithms become inoperative under high noise level and in the case of low channel bit error rate other very simple methods can be used.

Development of simple modification for block VA which, of course, belongs to methods with direct control of metrics, application of divergent decoding methods, MTD algorithms, classical and parallel methods of concatenation [5, 10, 11] create conditions for wider application of various coding methods widely represented in our web-sites www.mtdbest.ru and www.mtdbest.iki.rssi.ru. There also you can find some papers mentioned in the present work and other information, articles, presentations and books regarding the coding theory and its applied issues.

ACKNOWLEDGMENT

The research is carried out due to the support of the Ryazan State Radio Engineering University and Russian Foundation of Basic Research (grant 15-07-06348).

REFERENCES

[1] Viterbi A. Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm // IEEE Trans. 1967. IT-13. P.260-269.

[2] Johannesson R., Zigangirov K.Sh. Fundamentals of Convolutional Coding. Second Edition. John Wiley & Sons. 2015.

[3] Patent RF №2608872. The Method for Encoding and Decoding Block Code with Using of Viterbi Algorithm / Zolotarev V. V., Ovechkin P. V. (Space Research Institute of RAS). 2017. In Russian.

[4] Berrou C. Codes and Turbo Codes. Springer Science & Business Media, 2010.

[5] Zolotarev V.V., Zubarev Y.B., Ovechkin G.V. Optimization Coding Theory and Multithreshold Algorithms. Geneva, ITU, 2015. 159 p.

[6] Ma H.H., Wolf J.K. On Tail Biting Convolutional Codes // IEEE Trans. Commun. Feb. 1986. P.104-111.

[7] Kudryashov B.D. Basics of the Coding Theory. Saint-Petersburg: BKHV-Petersburg. 2016. 393 p. In Russian.

[8] Zolotarev V.V., Ovechkin G.V., Ovechkin P.V. About comparison of new error-correction methods // Proceeding of 18-th International Conference "Digital signal processing and its application – DSPA-2016". Moscow, 2016. Vol.1, P.59-64. In Russian.

[9] Arikan E. Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels // IEEE Transactions on Information Theory. 2009. Vol.55. No.7. P.3051-3073.

[10] Ovechkin G., Zolotarev V., Ovechkin P., Satibaldina D., Tashatov N. The Performance of Concatenated Schemes Based on Non-binary Multithreshold Decoders // Advances in Systems Science. Springer International Publishing, 2014. Vol.240, P.251–259.

[11] Zolotarev V.V., Ovechkin G.V. Divergent Encoding Convolutional Codes // Proceeding of 18-th International Conference "Problems of transmission and processing information in networks and telecommunication systems". Ryazan, Russia, 2015. P.27-32. In Russian.