

A LOW-COMPLEXITY SOFT VLC DECODER USING PERFORMANCE MODELING

Tsu-Ming Liu and Chen-Yi Lee

Institute of Electronics Engineering, National Chiao Tung University
1001, Ta Hsueh Road, Hsinchu, 300, Taiwan, R.O.C.

ABSTRACT

In this paper, we propose a scalable soft VLC decoder to greatly reduce overall design complexity. Generally, the soft VLC decoder needs to maintain many states for the correct decoding when the table size grows. We reduce the table size by using a symbol-merging scheme. We merge two symbols with the same prefix into one. Further, to achieve the optimal trade-off between performance and complexity, we propose a Black-Box model. In our model, we present a novel measurement of “symbol-alias” to improve the accuracy of performance estimation. Experimental results show that our scalable soft VLC decoder using performance modeling has more than 1dB PSNR gain and offers better subjective quality compared to traditional VLC decoding.

1. INTRODUCTION

Variable Length Codes (VLCs) are commonly used to approach the entropy rate of a given data source. They are extensively used in recent video coding standards. However, traditional table look-up hard decoding may lose synchronization and induce error propagation over a noisy channel. To improve the error resilience, the soft VLC decoders with joint source and channel design have been proposed [1]-[4]. Such algorithms generally need to maintain many states when the table size grows. Hence, soft VLC decoders have problems of high complexity.

Reduced complexity algorithms with sub-optimal solution have been made [5]. However, the improvement in [5] is not significant with larger VLC table. In this paper, we propose a scalable soft VLC decoder (VLD) to reduce the complexity. Firstly, our approach includes algorithm translation and table size reduction. We translate the metric derivation in Soft-Input Soft-Output (SISO) algorithm [6] into the symbol-constrained directed graph [5]. To reduce the table size, we propose a symbol-merging scheme. We merge two symbols with the same prefix into one symbol. We can greatly reduce table size as well as complexity at the cost of minor performance loss.

In [3], the authors used the minimal Hamming distance (d_H) to quantify the relation between table and performance. But, it is still inaccurate when the different tables reach the same d_H . We propose a novel measurement to improve the accuracy of performance estimation. Further, we reduce the penalty of over-design and observe the tendency of performance by the proposed Black-Box model. Thus, the proposed model reaches the optimal trade-off between performance and complexity.

*Work supported by the National Science Council of Taiwan, R.O.C., under Grant NSC 92-2220-E-009 -010-

The proposed scalable soft VLC decoder using performance modeling is verified with practical MPEG-4 table. From the analysis of performance and complexity, we can obtain the optimal parameters for a given table and decoding algorithm. Further, our proposed scheme shows more than 1dB PSNR improvement compared to the table look-up hard decoding.

The rest of this paper is organized as follows. Section 2 briefly introduces the SISO and presents our proposed symbol-merging scheme. Section 3 describes the proposed Black-Box model for the trade-off between performance and complexity. Section 4 presents the evaluation of MPEG-4 VLC table. Finally, section 5 summarizes our work and draws the conclusions.

2. SCALABLE SOFT VLC DECODER

2.1. Graph representation of SISO algorithm

SISO decoding technique [6] is considered as an exhaustive decoding procedure to resist the error disturbance in the noisy channel. It estimates and searches on the tree-like path in the existence of additive white Gaussian noise (AWGN). The input sequence is transmitted by packets. It uses L bits and equivalently N symbols to represent the priori information in one packet. Based on the similar formulation, we use the graph representation to simplify the SISO algorithm when the table size or decoded symbol grows.

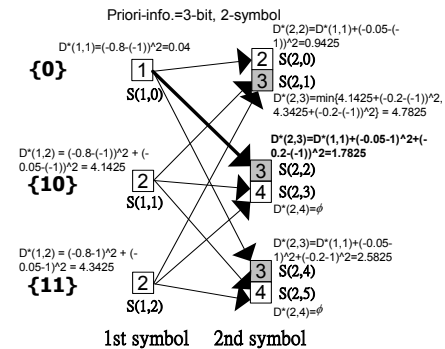


Figure 1. The proposed translation between the symbol-constrained directed graph and the SISO algorithm.

To help the understanding of our simplified algorithm, we use a symbol-constrained directed graph. Assume we have a simple VLC table with only 3 symbols {0,10,11} and a packet that includes 3 bits (equivalently 2 symbols) with content as ‘0 10’. After BPSK modulation, the modulated sequence is {-1, +1, -1}. When the packet is transmitted over the AWGN channel, the received packet may become {-0.8, -0.05, -0.2}.

Codelength	Codeword
1	0
2	10
2	11

(a) Simple VLC table

Codelength	Codeword	sign	LAST
1	0s	1	1
2	10s	1	0
2	11s	1	0

(b) Simplified MPEG VLC table

Figure 2. The original and real case of VLC table.

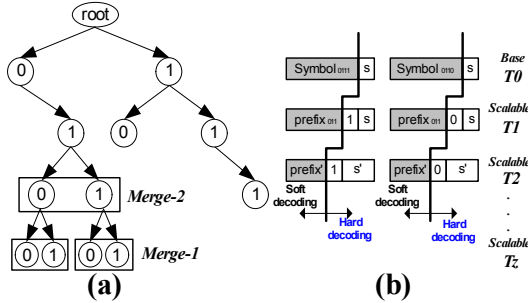


Figure 3. The scalable algorithm with hard and soft decoding.

Figure 1 depicts the graph representation for this example. The $D^*(i,j)$ denotes the cumulative square error of i -th symbol and equivalent j -th bit in each symbol-state. The $S(m,n)$ is the symbol state decoded with m -symbol and has the index of n . The operation of ‘minimum’ is exercised in the $S(2,1)$, which is entered by more than 2 arrows. Further, the minimal metric after the comparison is survived and the others are pruned. There is no need to calculate the $D^*(2,4)$ and $D^*(2,5)$ because the decoded bit pointer exceeds the priori information ($4 > 3$ bits). Therefore, we can decide the shaded squares as the final candidates. The $S(2,2)$ is the minimum, survives and traces back to the $S(1,0)$. Thus, we can decide the bitstream as $\{0,10\}$ for the correct decoding.

To apply our simplified algorithm in MPEG-4, we modify the simple VLC table as Figure 2(b). Specifically, each symbol is appended with an ‘s’ bit and each block is appended with a ‘LAST’ symbol to indicate the end of block decoding. The induced ‘sign’ field represents the number of ‘s’ in each symbol. Thus, after the modification of VLC table, we proposed the scalable scheme to reduce the table size in the following section.

2.2. Soft VLC decoder with symbol-merging algorithm

The main problem of soft VLC decoder is the many states and the complicated computation when the sequence length or table size grows. To apply the algorithm in MPEG-4, it is essential to reduce the table size. Thus, we propose a scalable scheme with symbol merging algorithm to reduce the decoding complexity.

We utilize the redundancy exhibiting in different symbols to perform the merging scheme. We consider the source table as a tree-structure in Figure 3(a). The symbol-merging scheme searches the symbols with identical prefix and merges them into the single symbol. In Figure 3(b), the original SISO algorithm is a special case, that is, when z is equal to 0 (T_0). In other words, there is no hard decoding performed except ‘s’ bits. Such case achieves the highest performance with the penalty of the largest complexity. However, the code-length of prefix symbol with soft decoding will decrease when the index ‘ z ’ increases. Meanwhile, the hard decoding bits will increase. Hence, it can be considered as a hybrid scheme that combines hard and soft decoding.

Code Length	Code Word	Sign	LAST
2	10s	1	1
3	111s	1	0
3	010s	2	0
3	011s	2	0

(a) $T_1 = 4$

Code Length	Code Word	Sign	LAST
2	10s	1	1
3	111s	1	0
3	01s	3	0

(b) $T_2 = 3$

Figure 4. A simple VLC table with ‘sign’ and ‘LAST’ field.

Table 1. The table size reduction by symbol-merging algorithm.

Standard	MPEG-2		MPEG-4	
Table	intra _{TB-15}	non-intra _{TB-14}	intra _{TB-16}	non-intra _{TB-17}
T_0	113	114	103	103
Scalable T_1	65	60	61	56
Scalable T_2	45	34	48	38
B.D.(2)	73.2%	90.7%	62%	77.8%

We use a simple example to illustrate the proposed scheme in Figure 4 where ‘ T_i ’ represents the number of symbols after the operation of $Merge-i$. As shown, after the operation of ‘Merge-1’, the table size is decreased by 2. Further, with the ‘Merge-2’ operation, the total number of symbols becomes 3. The introduced ‘sign’ field represents the number of ‘s’ appended in the corresponding symbol. The ‘sign’ field will increase when both of symbols have been merged into one.

To quantify the reduction capability after the symbol-merging scheme, we propose the ‘Balance Degree’ (B.D.) in Eq. (1). The metric of B.D. is between 0 and 1. In Eq. (1), the denominator represents the maximal value as well as a special table with complete tree-structure. It leads to ‘ $z \times 0.5$ ’ after the z times of summation where the ratio of T_{i+1} over T_i is fixed at 0.5. Therefore, the B.D. of Figure 4 is 58% in the condition of ‘ $z=2$ ’. To prove that B.D. is a meaningful number to our merging scheme, we measure the B.D. using the table in MPEG-2 and MPEG-4. As shown in Table 1, we find that the higher of the B.D., the more reduction of the table size. The B.D. of non-intra table is larger than that of intra one. It is explained by the fact that there is more redundancy exploited in the non-intra table.

$$B.D.(z) = \frac{\text{real reduction}}{\text{complete reduction}} = \frac{\sum_{i=0}^{z-1} \left(1 - \frac{T_{i+1}}{T_i}\right)}{\sum_{i=0}^{z-1} (1 - 0.5)} = \frac{\sum_{i=0}^{z-1} \left(1 - \frac{T_{i+1}}{T_i}\right)}{z \times 0.5} \quad (1)$$

3. THE PROPOSED BLACK-BOX MODEL

To optimize our proposed scalable soft VLC decoder, it is crucial to reach the optimal trade-off between performance and complexity. We introduce some parameters to describe the complexity. They are independent and composed of algorithm-sensitive, application-sensitive and table-sensitive. In the table-sensitive, we propose a novel measurement of ‘symbol-alias’ to provide accurate performance estimation for the different tables. Finally, the proposed Black-Box model can reach the optimal parameters for a given table and decoding algorithm.

Figure 5(a) depicts the proposed Black-Box (i.e. B-B) for the performance modeling and uses Figure 5(b) as the source VLC table for the following illustration.

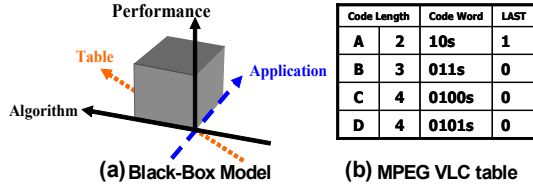


Figure 5. The B-B model and the evaluation of source table.

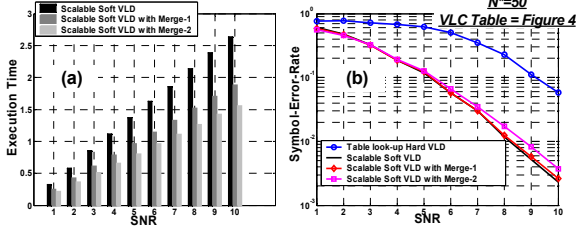


Figure 6. The complexity (a) and performance (b) in different ‘z’.

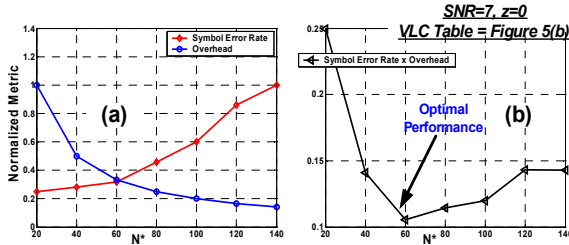


Figure 7. The optimization of performance in different ‘N*’.

3.1. Algorithm-sensitive: z

The algorithm-sensitive parameters are sensitive to algorithms, that is to say, different algorithms are characterized by different parameters. Using the proposed algorithm in section 2.2 as an example, the parameter ‘z’ is considered as an essential factor to approach the trade-off between performance and complexity. Figure 6 describes that the optimal choice is achieved when z is equal to 1 (Merge-1). The parameter ‘N*’ (see Section 3.2) does not affect the final results. Thus, ‘Merge-1’ obtains the maximal reduction of complexity at the cost of minor performance loss.

3.2. Application-sensitive: N

The application-sensitive parameters are unrelated to the algorithms and decided by the extrinsic applications. The packet size ‘N’ is the most impressive factor to achieve the optimal performance. In this section, we regard it as the decoded number of symbols ‘N*’ for the simplification. Figure 7(a) describes that the performance can be expressed by the normalized symbol error rate (i.e. SER) and overhead. In the overhead computation, we assume that the 15-bit resynchronization marker is inserted in the period of N*-symbol. From Figure 7(b), it achieves the minimal SER and overhead when ‘N*’ is equal to 60. There is no need to consider the algorithmic complexity because it has to be optimized from the algorithm-sensitive parameters.

3.3. Table-sensitive: T

Code Length	Code Word	
A	2	10
B	3	011
C	4	0100
D	4	0101

(a) Simple table in Figure 5(b)

Code Length	Code Word	
A	x_1	symbol ₁
B	x_2	symbol ₂
C	x_3	symbol ₃
D	x_4	symbol ₄

(b) Source table in Eq.(2)-(3)

$1/\text{Dist}_{\text{inter}} = 2 \rightarrow \{D,A,A\}$ or $\{C,A,A\}$ (i.e. $4=2+2$)

$\text{Dist}_{\text{intra}} = 5 \rightarrow 2 + 1 + 1 + 1$

Figure 8. The symbol alias of VLC table.

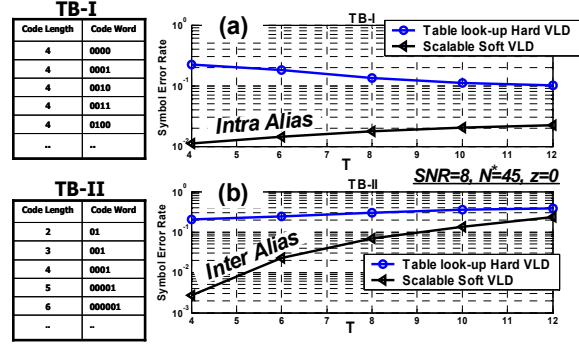


Figure 9. The performance evaluation in different ‘T’.

To estimate the performance with different tables, the authors in [3] used the minimal Hamming distance (d_H) to quantify the relation between VLC table and performance. It is still inaccurate when the different tables reach the same d_H . We propose a novel measurement of ‘symbol alias’ to quantify their relation and provide more accurate performance estimation when their Hamming distances are the same.

We introduce the table-sensitive parameters ‘T’ and the symbol alias to obtain more accurate estimation. The parameter ‘T’ (i.e. T_i in Fig. 4) denotes the number of symbols for a given VLC table. Besides, the symbol alias comprises two components. One is the distance of ‘Inter Alias’ (i.e. $\text{Dist}_{\text{inter}}$) described by Eq. (2). The other is the distance of ‘Intra Alias’ (i.e. $\text{Dist}_{\text{intra}}$) described by Eq. (3). Figure 8 depicts the results of the following equations. The $1/\text{Dist}_{\text{inter}}$ is the number of elements for the specific set, which calculates that whether the code-length of one symbol is the combination of the others. The $\text{Dist}_{\text{intra}}$ is the summation of minimal Hamming distance for each symbol.

$$\text{Set}_{\text{inter}} = \left\{ \{x_1, x_2, x_3, \dots, x_T\} \mid \sum_{i=1, x_i \neq \max\{x_i\}}^T x_i = \max\{x_i\} \right\} \quad (2)$$

$$\text{Dist}_{\text{inter}} = \frac{1}{\text{NumOfElement}(\text{Set}_{\text{inter}})}$$

$$\text{Dist}_{\text{intra}} = \sum_{i=1}^T \min(D_{i1}, D_{i2}, D_{i3}, \dots, D_{i, j \neq i}, \dots, D_{iT-1}, D_{iT}) \quad (3)$$

$$\text{where } D_{ij} = d_H\{\text{symbol}_i, \text{symbol}_j\}$$

The inter alias is more sensitive to ‘T’ than intra alias and induces more performance loss. In Figure 9(a), the intra alias of TB-I results from the bit alias with the symbols of identical code-length. The increase of ‘T’ provides the increase of SER in the soft decoding. However, the SER of table look-up decoding decreases when the ‘T’ grows. The reason is that the extra symbols prevent the decoded-symbol loss and error propagation. Further, the inter alias of TB-II results from the code-length alias with the symbols of different code-length. In Figure 9(b), both of SER increase and provide more performance loss than intra alias.

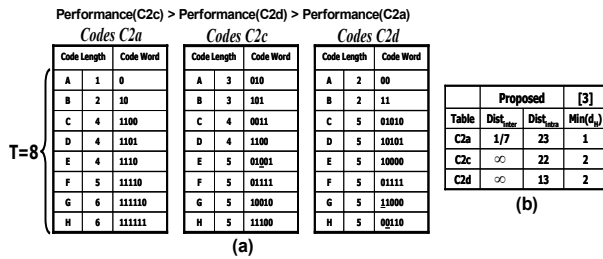


Figure 10. VLC tables and measurements for the same source.

To prove that the proposed symbol alias is a meaningful number, we use the VLC table in [3] to recognize the difference of performance. In Figure 10(a), we modify the underlined bits in C2c and C2d for the correct VLC decoding. For our proposed measurement of Figure 10(b), the Codes C2a achieves the worst performance because of the lowest $Dist_{inter}$ that induces more performance loss than $Dist_{intra}$. Further, there is no inter alias exist in C2c and C2d. The $Dist_{inter}$ is fixed at infinite (∞) by default. The C2c obtains higher performance than C2d because of the higher $Dist_{intra}$ of C2c. The relation of performance with different tables is identical to [3]. In Figure 10(b), the performance of C2c and C2d cannot be recognized in [3] when the occurrence of each symbol is unknown. We provide more accurate estimation than [3] and reduce the design time for the performance evaluation among different tables.

4. SIMULATION RESULT

Finally, we verify our proposed scalable soft VLC decoder (i.e. scalable soft VLD) over the AWGN channel using BPSK modulation. The input sequence is MPEG-4 encoded with the re-synchronization marker and data partition. In data partition mode, we have the texture part, composed of a sequence of VLC code-words, be corrupted by AWGN. The other parts are of error free. To the ESCAPE code, we simply use hard decoding.

To optimize the performance and the complexity, we include the B-B model in our evaluation of scalable soft VLD. We use foreman (QCIF) as our test sequence and encode the sequence at 64kbts/s and 25fps (No P-frame). In Figure 11(a), our proposed scalable soft VLD shows that more than 1dB PSNR can be gained compared to the table look-up decoding at $BER=10^{-3}$ (SNR=10dB). Further, the parameters determined by the B-B model are listed in Figure 11(b). The 'T' is determined from Table 1 with the given MPEG-4 table and the others are determined through our proposed B-B model. We choose 'z' as 2 for complexity reduction. Then, we obtain 'N' as 300 bits for performance optimization. The performance improvement of our proposed scheme will become more prominent when the upper bound of 'No error' is increased. In the subjective quality comparison of Figure 12, our scheme shows better quality.

5. CONCLUSION

In this paper, we present a scalable scheme of soft VLC decoder to reduce overall design complexity. By the symbol-merging scheme, we can greatly reduce the table size with the price of minor performance loss. Further, we propose a novel measurement of 'symbol-alias'. We also provide more accurate

performance estimation than [3] for the different tables. With the proposed B-B model, we can achieve the optimal trade-off between performance and complexity. For the proposed scalable soft VLC decoder using performance modeling, we averagely improve the PSNR by 1dB and offer better subjective quality compared to the table-look-up hard decoding.

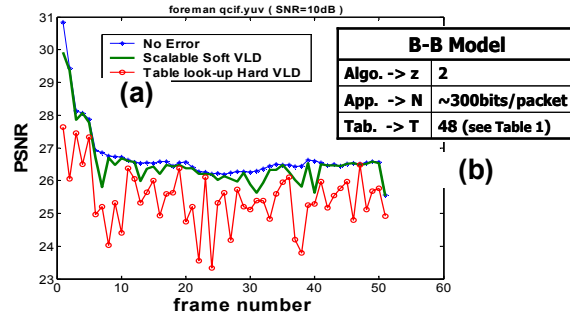


Figure 11. PSNR comparison within 50 frames.

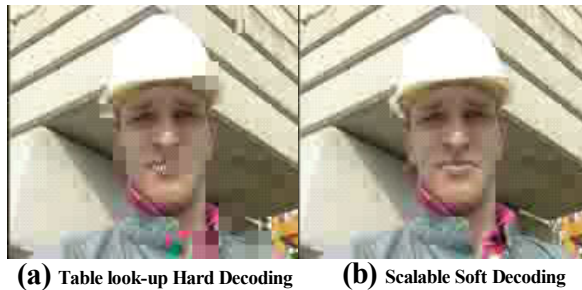


Figure 12. The comparison on the 1st frame of video sequence.

6. REFERENCES

- [1] Jeanne, M.; Carlach, J.C.; Siohan, P.; Guivarch, L.; "Source and Joint Source-Channel Decoding of Variable Length Codes" *IEEE International Conference on Communications*, vol. 2, pp.768-772, May 2002.
- [2] Demir, N.; Sayood, K., "Joint source/channel coding for variable length codes", *Data Compression Conference '98, Proceedings*, pp. 139-148, April 1998.
- [3] Bystrom, M.; Kaiser, S.; Kopansky, A., "Soft source decoding with applications", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, Issue. 10, pp. 1108 - 1120, Oct. 2001.
- [4] M. Park and D. J. Miller, "Joint source-channel decoding for variable length encoded data by exact and approximated MAP sequence estimation", *ICASSP '99*, Phoenix, Arizona, USA, pp. 2451-2454, March 15-19, 1999.
- [5] Lamy, C.; Pothier O., "Reduced complexity Maximum A Posteriori decoding of variable-length codes", *GLOBECOM '01*, vol.2, pp.25-29, Nov. 2001.
- [6] Jiangtao Wen; Villasenor, J., "Soft-input soft-output decoding of variable length codes", *IEEE Transactions on Communications*, Vol. 50, Issue: 5, pp. 689-692, May 2002