**IV. THE PROPOSED FAST INTRA PREDICTION PROCESS**

In order to improve the coding speed, we have conducted extensive experiments to investigate of Intra coding in spatial SHVC. In order to ensure generality of the features and rules, different sequences, including motion and texture from simple to complex, are selected. In this research, we use Blue-sky, Ducks, Park\_Joy, Pedestrian, Tractor, Town and Station2 in testing. According to common SHM test conditions (CSTC) [27], there exist scalability ratio 2x and 1.5x in spatial SHVC. In 2x, ratio of both height and width in EL to those in BL is 2; in 1.5x, ratio of both height and width in EL to those in BL is 1.5. Each scalability ratio also includes one set QP in BL and two set QP setting in EL: the QPs in the BL are (22, 26, 30, 34), and the corresponding QPs in the EL are (22, 26, 30, 34) and (24, 28, 32, 36), respectively. Since there are two scalability ratios, and each ratio includes two QPs, there are four sets of scalability ratios and QPs. Among the four sets of scalability ratio and QPs, 2x and QPs (24, 28, 32, 36) in EL should have the weakest inter-layer correlation. If it can achieve good performance, the other QPs set can obtain even better performances. Therefore, we only use 2x and QPs (24, 28, 32, 36) in EL in conducting experiments. Based on these experiments, we propose the corresponding fast Intra prediction strategies below.

1. **Residual Coefficients of ILR Mode-Based Depth Early Skip (RCIM-BDES)**

In SHVC, each CTU includes four depths, corresponding to sizes of a CU from 64×64 to 8×8. Each CU needs to check both ILR mode and Intra mode, and then select the mode with the smaller rate distortion (RD) cost as the best mode. In order to investigate mode distribution between ILR mode and Intra mode, we use the aforementioned condition in testing. The corresponding mode distribution between ILR mode and Intra mode is listed in Table I.

Table I mode distribution between ILR mode and Intra mode

|  |  |  |
| --- | --- | --- |
| Sequence | ILR | Intra |
| Blue-sky | 98.50% | 1.50% |
| Ducks | 99.76% | 0.24% |
| Park\_Joy | 99.07% | 0.93% |
| Pedestrian | 90.65% | 9.35% |
| Tractor | 97.68% | 2.32% |
| Town | 96.84% | 3.16% |
| Station2 | 95.27% | 4.73% |
| Average | 96.82% | 3.18% |

From Table I, we can find that the average percentage of ILR mode is 96.82%. In other words, most CUs select ILR mode as the best mode. The reason is that the content in BL and EL are the same, and QPs in BL and EL are similar or even the same, the inter-layer correlation is very strong. In addition, the prediction of ILR mode is obtained directly by upsampling the co-located CUs in BL, so the process is very simple. Therefore, we can first obtain the residual coefficients of ILR mode without encoding it, and determine whether the current CU needs to be further split based on the residual coefficients. In the affirmative, the current CU can be directly skipped. Otherwise, we need to further encode ILR mode and Intra mode.



Fig.3. The division of the CU.

We first divide the CU of the residual coefficients into top and bottom parts as shown in in Fig. 3 (a) and left and right parts as shown in in Fig. 3 (b). Obviously, if two parts of any division are significantly different, it means that the current CU needs to be further split. In general, if a CU is predicted very well by the best mode, the corresponding residual coefficients will follow Gauss distribution [18]. Suppose the residual coefficients obey a Gaussian distribution, the residual coefficients of one part in each division are respectively modeled as:

 (1)

whereis the residual coefficients of the part, andare respectively its expected value and variance. Supposeare samples in, theandcan be derived by maximum likelihood estimation (MLE). The probability density function ofis:

, (2)

its corresponding likelihood function is:

, (3)

 , (4)

where is the number of residual coefficients in each part. In order to obtain and, we can calculate below:

 (5)

From Equ. (5), we can deriveandcan be obtained and they are:

 (6)

Through the above process, we can obtainandof one part, and then we use them to test whether the residual coefficients of the other part also use them. Supposeis the residual coefficients of the other part, are its samples. We can test whetheralso useandas follows:

, (7)

whereis the significance level value. For any, the corresponding threshold valuecan be obtained by checking Gaussian distribution table. If (7) is satisfied, the residual coefficients of the two parts do not use the same expected value and variance. Therefore, the two parts are significantly different, the current CU needs not to be checked and be skipped directly.

Since different depths may have different probability to be skipped, so we need to select the best threshold value of every depth. For depth 2, we select some common values to check, their corresponding coding efficiencies are relatively large. In order to improve coding efficiency, we select the largest values 3.49 in Gaussian distribution table, and further use its multiples in testing, the corresponding coding efficiency are listed in Table 2.

Table 2. coding efficiency under different test values

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Test values  Sequence | 3.49 | 6.98 | 10.47 | 13.96 | 17.45 | 20.94 | 24.43 | 27.92 |
| blue\_sky | 0.36% | 0.40% | 0.38% | 0.27% | 0.15% | 0.10% | 0.04% | 0.04% |
| ducks | 0.10% | 0.09% | 0.08% | 0.03% | 0.01% | 0.00% | 0.00% | 0.00% |
| park\_joy | 0.43% | 0.34% | 0.20% | 0.09% | 0.04% | 0.01% | 0.01% | 0.00% |
| pedestrian | 0.17% | 0.25% | 0.23% | 0.17% | 0.18% | 0.09% | 0.04% | 0.00% |
| town | 0.32% | 0.29% | 0.23% | 0.12% | 0.04% | 0.02% | 0.01% | 0.00% |
| station2 | 0.00% | 0.07% | 0.12% | 0.08% | 0.04% | 0.05% | 0.02% | 0.01% |
| tractor | 0.13% | 0.14% | 0.13% | 0.10% | 0.04% | 0.02% | 0.00% | -0.01% |
| Average | 0.22% | 0.23% | 0.20% | 0.12% | 0.07% | 0.04% | 0.02% | 0.01% |

In Table 2, coding efficiency are denoted by BDBR [28], which measures the bitrate difference at equal PSNR in the EL. A positive or negative BDBR reflects a coding efficiency loss or increase, respectively. We can observe that when test value is greater than or equal to 20.94, BDBRs in all sequences are less than 0.1%. Therefore, test value 20.94 is selected as threshold value in depth 2. In the similar way, we obtain threshold values for depth 1 is 31.41. If we skip depth 0, the corresponding coding efficiencies will degrade obviously in some sequences. Therefore, we do not skip depth 0. Based on the above analysis, we can rewrite the depth skip conditions as follows:

, (8)

When the above condition is satisfied, we can directly skip the corresponding depth.

1. **RD Cost of ILR Mode-Based Mode Selection (RCIM-BMS)**

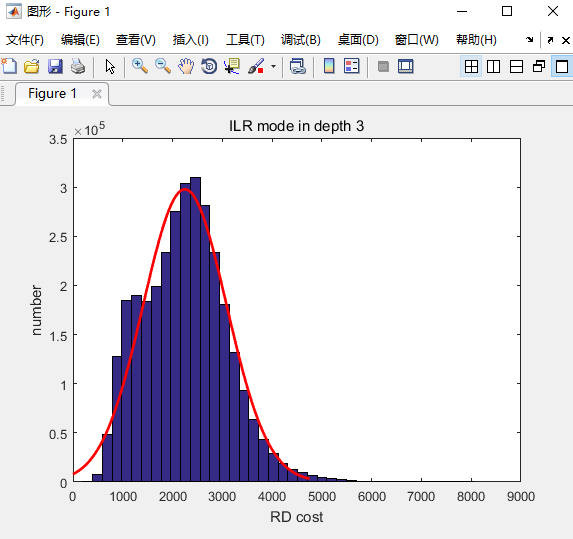
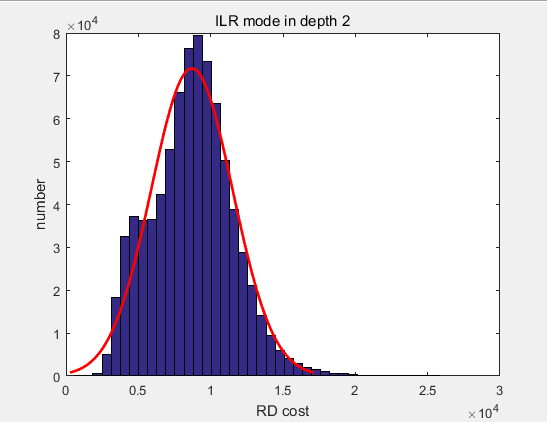
As mentioned above, ILR mode occupies the majority of mode distribution. However, in some sequences, Intra mode also occupies nonnegligible proportion, such as sequence “Pedestrian”. If we only use ILR mode to check, the corresponding coding efficiency may be obviously degraded in some sequence. However, if we use both ILR mode and Intra mode to check, much unnecessary coding time would be costed in many sequences. In order to improve the coding speed and maintain the coding efficiency, we need to investigate mode selection principle behind. Using the aforementioned condition in testing, we can obtain RD Costs of ILR mode and Intra mode, which are list in Table 3.

Table 3. RD Cost of ILR mode and Intra mode

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sequence | Depth 0 | | Depth 1 | | Depth 2 | | Depth 3 | |
| ILR | Intra | ILR | Intra | ILR | Intra | ILR | Intra |
| Blue-sky | 49801 | 11074 | 12172 | 3198 | 3039 | 1523 | 778 | 946 |
| Ducks | 140109 | 49873 | 34478 | 13686 | 8701 | 5732 | 2247 | 2444 |
| Park\_Joy | 166607 | 21382 | 41191 | 5616 | 10308 | 3066 | 1888 | 2621 |
| Pedestrian | 37458 | 26252 | 9333 | 7063 | 2351 | 2189 | 609 | 763 |
| Tractor | 57344 | 22816 | 14445 | 5734 | 3719 | 2433 | 988 | 1189 |
| town | 108752 | 63737 | 27495 | 16467 | 6875 | 4908 | 1721 | 1981 |
| station2 | 44750 | 12625 | 11201 | 4611 | 2839 | 2300 | 742 | 901 |
| average | 86403 | 29680 | 21474 | 8054 | 5405 | 3164 | 1282 | 1549 |

From Table 3, we can observe that RD Cost of ILR mode is significantly larger than that of Intra mode from depth 0 to 2; while RD Cost of ILR mode is smaller than that of Intra mode in depth 3. The reason is ILR and Intra mode use different processes in prediction. ILR mode use co-located pixels in BL to upsample in horizontal and vertical directions. Intra mode uses 35 directional modes in prediction. For large CUs, namely depth from 0 to 2, reference pixels and pixels in the current CU correlation is not very strong. If texture is very complex, co-located pixels in BL can predict the current CU better than reference pixels, so ILR mode is selected as the best mode. If texture is very simple, 35 directional modes can predict the current CU better than upsampling in horizontal and vertical directions, so Intra mode is selected as the best mode. For small CUs, namely depth 3, reference pixels and pixels in the current CU correlation is very strong. If texture is very complex, Intra mode with 35 DMs can predict the current CU better than ILR mode, so Intra mode is selected as the best mode. If texture is very simple, both ILR mode and Intra mode can predict very well, the difference of their RD cost is also very small, so either of them can be selected as the best mode. Therefore, the average difference of their RD cost is not very significantly.

In addition to the difference of RD Costs of ILR mode and Intra mode, we further investigate the RD cost distribution of ILR mode and Intra mode. In Fig.2, the horizontal axis represents RD costs, and the vertical axis represents the corresponding numbers. Fig.2. (a) and Fig.2. (b) shows that the RD cost distribution of ILR mode with sequence “Ducks” in depth 2 and depth 3, respectively.



(a) (b)

Fig.4. The RD cost distribution of ILR with sequence “Ducks”

From Fig.4, we can observe that both the RD cost of ILR in depth 2 and depth 3 follow Gaussian distribution. Extensive experiments show that RD cost distribution of ILR mode and Intra mode in all sequences follow Gaussian distribution.

Based on the above analysis, we can draw a conclusion: RD Cost of ILR mode is different from that of Intra mode in all depths, and their distributions all follow Gaussian distribution. Based on this feature, we propose to adopt Gaussian Mixture Model and Expectation Maximization (GMM-EM) to determine the best mode. Since ILR mode occupies the majority of mode distribution, we can encode ILR mode, and then use GMM-EM to determine whether ILR mode is the best mode based on its RD cost. In the affirmative, we can directly skip Intra mode; otherwise, we need to further check Intra mode.



Fig.5. The current CU and its Relative CUs

In order to use GMM-EM in prediction, we use coding information modes and RD costs of relative CUs of the current CU in prediction. As shown in Fig.5, is the current CU, ,,andare the neighboring of the current CU, ,,, andare the co-located CUs of ,,,and. For any , its corresponding RD cost is denoted as. The corresponding Gaussian Mixture Model is:

, (9)

whereis the possibility of using ILR mode,andare respectively expected value and variance of their RD cost; andis the possibility of using Intra mode,andare respectively expected value and variance of their RD cost. *M* is the number of the current CU and its relative CUs, and it is 10.

In order to obtain these six parameter values, the maximum likelihood estimation is used in calculation as follows:

 (10)

The logarithm of the likelihood function is:

 (11)

,,can been calculated by:

  (12)

We can derive:

 (13)

where, then we have：

 (14)

is the probability thatis generated by the *k-th* part, and it can be obtained by：

 (15)

Repeat iterations (13), (14) and (15) untilconverges.

Since the current CU is, in order to decide whether ILR mode is the best mode, we need to determine if converges. Suppose the i-th iteration ofis denoted as. In order to avoid unnecessary repeat iterations, if the absolute difference betweenandis very small, we can terminate the repeat iteration. We can empirically select 0.01 as a threshold, then we have:

. (16)

If condition (16) is met, we can terminate the repeat iteration. Through the above process, we can obtain the possibility of the current CU selecting ILR mode. Since this possibility is obtained based on *RD* cost, we define it as *RD* based probability.

Since relative CUs are usually very similar, we can use relative CUs in prediction. Obviously, if more relative CUs use ILR mode, the current CU is more likely to use this mode, vice versa. In other words, the probability of current CU selecting ILR mode is proportional to the number of relative CUs using ILR mode. As shown in Fig.3, the current CU has nine relative CUs. Therefore, the possibility of the current CU selecting ILR mode can be written as , whereis the number of relative CUs using ILR mode. Since this possibility is obtained based on the number of relative CUs using ILR mode, we define it as number-based probability.

Since both *RD* based probability and number-based probability have strong relationship with the ILR mode selection, we can combine with them to predict the probability to use ILR mode. Let A and B denote the *RD* based probability and the number-based probability, respectively. Obviously, both of them are independent, we can derive:

 (17)

wheredenotes the possibility of depth early termination. Ifis greater than or equal to 0.6, the current CU is very likely to be early terminated. Therefore, we use 0.6, 0.7, 0.8 and 0.9 during testing, and the corresponding BDBRs are presented in Table 4.

Table 4. The possibility of depth early terminationand the corresponding BDBRs

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Pr & BDBR | 0.6 | 0.7 | 0.8 | 0.9 |
| Blue-sky | -0.3% | -0.3% | -0.3% | -0.3% |
| Ducks | 0.0% | 0.0% | 0.0% | 0.0% |
| Park\_Joy | 0.0% | 0.0% | 0.0% | 0.0% |
| Pedestrian | -0.1% | -0.1% | -0.1% | -0.1% |
| Tractor | -0.1% | 0.0% | 0.0% | 0.0% |
| town | -0.1% | -0.1% | -0.1% | -0.1% |
| station2 | 0.0% | -0.1% | -0.1% | -0.2% |

From Table 4, we can find that BDBR stays the same except in sequence “station2” under different. Whenis equal to 0.9, BDBR achieve the smallest value in sequence “station2”. Therefore, we select 0.9 as the best value for.

1. **DM Distribution-Based DM Selection (DD-BDS)**

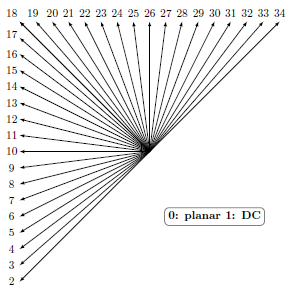


Fig.6. 35 DMs in SHVC

Similar to HEVC, SHVC also use 35 DMs in prediction, as shown in Fig.6. Through checking the 35 DMs, the DM with the smallest HC can be obtained which is defined as the best DM. As mentioned in the above section, for large CUs, simple CUs usually use Intra mode; for small CUs, complex CUs usually use Intra mode. In order to obtain their DM distributions, we use the aforementioned condition in testing. The corresponding DM distributions in large CUs and small CUs are list in Table 5.

Table 5. DM distribution in large CUs and small CUs

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sequence | Large CU | | Small CU | |
| 0 | 1 | 0 | 3 |
| Blue-sky | 68.23% | 82.38% | 76.41% | 83.93% |
| Ducks | 20.28% | 90.78% | 30.37% | 85.39% |
| Park\_Joy | 66.55% | 83.84% | 67.70% | 82.39% |
| Pedestrian | 57.92% | 93.08% | 69.89% | 94.62% |
| Tractor | 38.31% | 73.00% | 52.58% | 81.85% |
| Town | 72.85% | 94.80% | 61.73% | 92.59% |
| Station2 | 41.27% | 70.45% | 54.74% | 74.32% |
| Average | 52.20% | 84.05% | 59.06% | 85.01% |

In Table 5, 0 and 1 refer to class 0 and 1. Class 0 refers to DM 0 and 1, Class 1 refers to DM 0, 1, 8, 9, 10, 11, 12, 24, 25, 26, 27 and 28. From Table 6, we can find that more than 50% CUs use class 0, 85% CUs use class 1. The reason is as follows: if large CUs use Intra mode, they are usually very simple, so the distribution of their DMs is very regularly; while for small CUs, their sizes are very small so their texture cannot change very significantly, so the distribution of their DMs also are very regularly. Since class 1 occupies about 85%, the other DMs occupy about 15%. We define the other DMs as class 2.

Although class 0 only includes 2 DMs, it occupies more than 50% possibility. If we can determine the best DM in class 0, we can significantly improve the coding speed. Similarly, class 1 includes 12 DMs, it occupies about 85% possibility. It means that the best DM is very likely in class 1. If HC of a DM is smaller than those of its neighboring two DMs, we define the DM as local minim DM(LMD). The best DM should be identified within LMDs. After checking class 1, if there is a LMD, it is very likely to be the best DM, so we can early terminate DM selection. Otherwise, we need to further check class 2. Since the probability of class 2 is very small but cannot totally ignored. If we always check all the DMs, it will cost much unnecessary time. However, if we directly skip these DMs, the coding efficiency will be obviously degraded. Therefore, we use variable step and binary search to obtain the best DMs. For the convenience of later description, we defineas the HC of DM. Based on the above analysis, we propose the corresponding method to predict the candidate DMs as follows:

1. We first select DM 0, 1, 10 and 26 to check. The smaller HC in DM 0 and 1 is denoted as min (0, 1), the smaller HC in DM 10 and 26 is denoted as min (10, 26). If min (0, 1) is significantly smaller than min (10, 26), DM 0 and 1 are very likely to be the best DM, go to (10); else go to (2).
2. We comparewithto further select likely DMs: ifis significantly smaller than, go to (3); else ifis significantly smaller than, go to (5); else go to (7).
3. The best DM is very likely to in horizontal DMs. Since the best DM is very likely in class 1, we further check DM 8, 9, 11, 12. If there is a LMD within DM 9, 10 and 11, the DM is very likely to be the best DM, go to (10). Otherwise, go to (4).
4. we further check DM 2, 6, 14 and 18. If the DM with the smallest HC is not within DM 2, 6, 8, 12, 14 or 18, the best DM has a very small probability in class 2, there is no need to further check other DMs in class 2, go to (10). Otherwise, binary search is used to obtain the best DMs, go to (9).
5. The best DM is very likely to in vertical DMs. Since the best DM is very likely in class 1, we further check DM 24, 25, 27, 28. If there is a LMD within DM 25, 26 and 27, the DM is very likely to be the best DM, go to (10). Otherwise, go to (6).
6. We further check DM 18, 22, 30 and 34. If the DM with the smallest HC is not within DM 18, 22, 24, 28, 30 and 34, the best DM has a very small probability in class 2, there is no need to further check other DMs in class 2, go to (10). Otherwise, binary search is used to obtain the best DMs, go to (9).
7. Since the best DM is very likely in class 1 and DM 10 and 26 have already been checked, we further check class 2 except DM 10 and 26. If there is a LMD within DM 9, 10, 11, 25, 26 and 27, the DM is very likely to be the best DM, go to (10). Otherwise, go to (8).
8. We further check DM 2, 6, 14, 18, 22, 30 and 34. If the DM with the smallest HC is not within DM 2, 6, 8, 12, 14, 18, 22, 24, 28, 30 and 34, there is no need to further check other DMs in class 2, go to (10). Otherwise, binary search is used to obtain the best DMs, go to (9).
9. The best DM is very likely within class 2. We further check the middle of the DM with the smallest HC and its left (right) checked neighboring DMs and select the DM with the smallest HC, repeat the process until a DM is a LMD. The DM is very likely the best DM, go to (10).
10. The DM selection is terminated.

In order to decide whether two DMs’ HCs have significant difference, their corresponding residual coefficients are investigated. Suppose and  are residuals of two DMs, their differenceis:

 (18)

Through Hadamard transformation, the above equation can be rewritten as:

, (19)

where *H* is aHadamard matrix. According to Cauchy-inequality, we can derive:

 (20)

where *m* is the size of the current CU. Then

 (21)

is a value in *HRH* at (*i, j*) position, which is calculated by:

 (22)

Ifany quantized values in *HRH* is smaller than, andare not significantly different. The following condition should be satisfied:

 (23)

Combine Eq. (18), (19), (20), (23), we can derive:

 (24)

Eq. (24) can be written as:

, (25)

whereandrefers to Hadamard transform values of two IMs. If Eq. (25) is satisfied, we can consider that they are not significant different. Conversely, the significant different condition can be written as:

, (26)

If Eq. (26) is satisfied, they can be considered to be significant different. In order to obtain the best , we use the aforementioned condition in testing. The corresponding the corresponding BDBRs are list in Table 6.

Table 6. and the corresponding BDBRs

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 |
| blue\_sky | 0.00% | 0.00% | 0.00% | 0.01% | 0.01% | 0.00% |
| ducks | -0.01% | -0.01% | 0.00% | -0.01% | 0.00% | 0.00% |
| park\_joy | -0.01% | -0.03% | 0.00% | 0.00% | 0.01% | 0.01% |
| pedestrian | -0.02% | 0.00% | 0.00% | 0.00% | 0.00% | 0.01% |
| town | -0.16% | -0.11% | -0.10% | -0.07% | -0.03% | -0.02% |
| station2 | 0.16% | 0.14% | 0.13% | 0.11% | 0.08% | 0.06% |
| tractor | 0.01% | 0.00% | 0.01% | 0.00% | -0.01% | -0.01% |

From Table 6, we can find that there is a turning point whenis equal to 5. If greater than or equal to 5, the corresponding BDBRs in all sequences are small than 0.1%. It means that *k* with 5 can obtain very good performance. If we further select larger, the corresponding coding speed improves will be smaller. Therefore, is set to 5.

In order to describe conveniently, we use “<<” to represent significantly smaller than. The corresponding flowchart is shown in Fig.6.



Fig.6. Flowchart of DM selection

1. **Residual Coefficients of Depth-Based Depth Early Termination (RCD-BDET)**

After the current CU has been checked, we can obtain its residual coefficients. Similar to depth skip, we also use two ways to divide the CU of the residual coefficients into two parts, as shown in Fig.3. If residual coefficients of two parts in each division do not have significant difference, the current depth is very likely to be the best depth. Therefore, the current CU needs not to be further split and depth selection can be early terminated. Also similar to depth skip, we first use MLE to calculate theandof one part, and then use them to test whether the other parts also use them. Supposeis the residual coefficients of the other part, are its samples. We can test whetheralso useandas follows:

, (27)

whereis the significance level value, is the number of residual coefficients in each part. By checking Gaussian distribution table, we can obtain the corresponding threshold value. If (7) is satisfied, the residual coefficients of the two parts use the same expected value and variance. Therefore, the two parts do not have significant difference, the current CU can be early terminated.

We select some commonvalues to check, such as 0.0005, 0.0015, 0.0025, 0.0045, 0.0125, 0.015 and 0.025. The corresponding threshold value are 3.3, 2.96, 2.81, 2.61, 2.24, 2.17 and 1.96, respectively. Experiments shown their corresponding coding efficiency losses are relatively large. In order to improve, we first divide 1.96 by 2, and then divide the value by 2, repeat the process until coding efficiency losses are very small. Based on experiments, the corresponding thresholds for depth early termination are list below:

 (28)