Journal of Applied Science and Engineering, Vol. 19, No. 2, pp. 229–239 (2016) *DOI: 10.6180/jase.2016.19.2.14*

Estimating Rate-Distortion for Multiple Description Coding Based on Zero Padding

Rong-Hua Chang, Chieh-Wei Su and Chow-Sing Lin*

Deparment of Computer Science and Information Engineering, National University of Tainan, Tainan, Taiwan 700, R.O.C.

Abstract

In the studies related to multiple description coding (MDC), applying zero padding not only improves the spatial correlation between descriptions but also increases their sizes. However, as we pad more zeros, the spatial correlation does not improve linearly. Therefore, the distortion is not reduced as the same amount when the size increases. In this paper, we focus on how to efficiently find an acceptable trade-off between the number of padded zeros and the increase of distortion in MDC. In this paper, we propose two distortion estimations for MDC with zero padding based on spatial quincunx distribution, which are standard deviation for distortion estimation and triangular linear approximation distortion estimation (TLA-DE) to efficiently calculate the distortion for the number of padded zeros atthe horizontal direction. In order to search for the acceptable number of padded zeros, we propose two search methods, Naïve Search and Best Division Search, which are both based on the proposed rate-distortion trade-off indicator. Our experiments result show that the worst error of distortion estimation by TLA-DE is up to 18%, but the average is 5.5% with 95% confidence level. The executing time of best division search is about two times faster than that of naïve search for 11 test sequences. To sum up, we confirm that the acceptable trade-off between rate and distortion in low computational complexity can be found by using the best division search on TLA-DE.

Key Words**:** Multiple Description Coding, Zero Padding, Rate-Distortion, Scalable Video Coding

1. Introduction

Adaptive coding has recently been introduced to address the variety of device capability on heterogeneous network for providing multimedia services. One of major techniques for adaptive coding is the multiple description coding (MDC). MDC could encode a single media sequence, such as video or audio, into equal- or unequal-rate multiple sub-streams, which are called descriptions [1–3]. Each description is independent and self-decodable. Any number of descriptions can be arbitrary combined and decoded to provide partial media quality commensurate with the network, display, and computing capability of a client [4–7].

In general, a sequence can be divided into multiple descriptions by MDC based on three domains: spatial [8–11], temporal [12,13], and frequency [14]. In this paper, we limit the discussion to divide a sequence into multiple descriptions on spatial domain for MDC. Generally, in MDC descriptions generated by dividing a sequence in spatial domain exist high spatial correlation. In order to further increase the spatial correlation to enhance the video reconstruction quality in a side decoder, we may add extra data, such as padding zeros in the frequency domain transformed by discrete cosine transform (DCT) to control the redundancy of a video. Zero padding in the DCT domain means oversampling in the spatial domain, in which a certain amount of redundancy are added to the image. By padding zeros in high frequency during DCT, the error concealment of MDC would be significantly improved [15]. Recently, MDC with zero padding has also been applied to network coding [16] and video encryption [17].

To develop multiple description coding with zero

*Corresponding author. E-mail: mikelin@mail.nutn.edu.tw

padding based on spatial Quincunx distribution (QZP-MDC) in this paper, we derived two distortion estimations of QZP-MDC with two descriptions, called Standard Deviation for Distortion Estimation (STD-DE) and Triangular Linear Approximation Distortion Estimation (TLA-DE). We also designed the Rate-Distortion Trade-off Indicator (RDTI) to measure the trade-off between distortion and the number of padding zeros. We also propose Naïve Searchand Best Division Searchto quickly output the result once an acceptable RDTI is obtained since the time complexity of computing RDTIs for all possible number of padding zeros is proportional to the width of an image. Our experiment results show thatthe worst error of distortion estimation by TLA-DE is up to 18%, but the average is 5.5% with 95% confidence level. The executing time by best division search is about two times faster than that of naïve search for 11 test sequences. To sum up, we confirm that the acceptable trade-off between rate and distortion in low computational complexity can be found by using the best division search on TLA-DE.

The remainder of this paper is organized as follows. In section 2, we briefly the concept of describe zero padding. Multiple description coding with zero padding based on spatial quincunx distributionis introduced in section 3. Proposed approaches for distortion estimation in zero padding are described in sections 4 and 5. In section 6, we describe how to find the acceptable trade-off between rate and distortion. In section 7, we describe experimental results. Finally, the summary of our work in presented in section 8.

2. Zero Padding

When a description is lost,typically we can utilize error concealment (EC) [15,18,19] to fix missing information by the spatial correlation among pixels. In general, improving the spatial correlation between descriptions can further increase the EC efficacy. Padding zeros in time domain can result in interpolation in frequency domain. On the other hand, padding zeros in frequency domain results in interpolation in time domain [20–22].

Wang et al. proposed one dimension of MDC (1D-MDC) [14,17] based on1-D of DCT (1D-DCT). This technique was considered for image encoding and it provided sound results. In 1D-MDC encoding, first each frame will be transformed by using 1D-DCT on each column and then padding extra zeros beyond highfrequency. Next, After 1-D of inversed DCT (1D-IDCT), extra columns are added into the original frame which is then sub-sampled into two sub frames and each are independently coded as a description, as depicted in Figure 1(a). After zero padding, pixels are more spatially correlated, and they can still approximately restorethe original quality by EC even one description is lost. When only one description is received at a side decoder, as in Figure 1(b), the other description will be estimated based on the received descriptions. They can be then merged together to conceal errors and fix missing information. Next, the padded zeros are removed once the merged frame is transformed to frequency domain by 1D-DCT. Finally, the 1D-IDCT is applied to get the original frame. How-

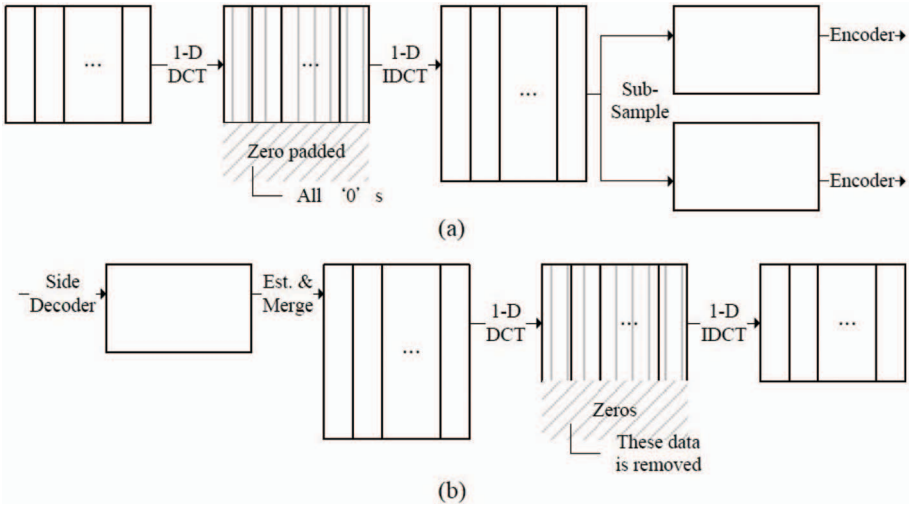


Figure 1. Encoder and Decoder by 1D-MDC with zero padding [14].

ever, this technique pads the same number of zeros for every frame without considering the complexity of the frames. When most objects in a frame are smooth or the content of a frame is simple, excessively padding zeros cannot further improve spatial correlation, resulting in additional computation. Therefore, the efficacy of zero padding can be further improved by preprocessing each frame to decide how many zeros should be padded. Such a concept is applied in our paper to try to find the best number of padded zeros in each frame.

3. Multiple Description Coding with Zero Padding Based on Quincunx Pixel Distribution

In this section, we introduce a basic MDC method, named multiple description coding with zero padding based on spatial quincunx distribution (QZP-MDC), for analyzing the trade-off between the rate and distortion. Figure 2 shows the data flow for multiple description coding with quincunx pixel distribution based on zero padding. First, each frame in the sequence is extended by padding zeros at frequency domain to increase the spatial correlation. The resulted spatial pixels then are encoded by quincunx distribution into two descriptions which can be sent to error-prone networks. Any errors can be concealed to reconstruct the sequence by means of nearby pixels.

As shown in Figure 3, the pixels in those frames are distributed by quincunx mapping as two descriptions. Quincunx mapping is like a chess board [15]. A quincunx map contain white blocks, Quincunx Signal A (QSA), and the non-white blocks, Quincunx Signal B (QSB). Each pixel in QSA and QSB, except for the pixels locating at edges and corners, is surrounded by the pixels in the other pixels group. Therefore, we can efficiently utilize the means of nearby pixels to conceal missing pixels.

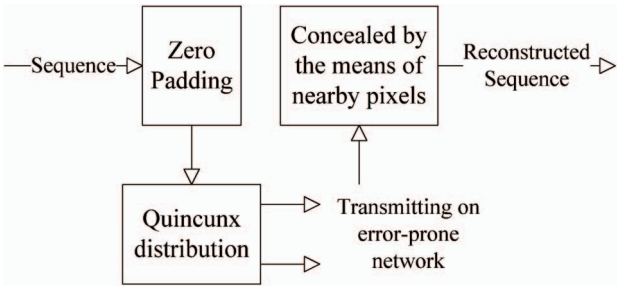


Figure 2. The data flow for multiple description coding with quincunx pixel distribution based on zero padding.

While transmitting descriptions on error-prone network, some packets may be dropped due to unreliable channels. In order to avoid quality deterioration, a lost description is concealed by nearby pixels from the other description. As Figure 3 shows, the missing pixel X in the missing description QSA is replaced by the mean of a, b, c, and d.

4. Distortion Estimation in Zero Padding

Before we discuss the rate and distortion, we want to obtain the actual distortion of each frame with different numbers of padded zeros. It would take too much time to calculate every frame, so we propose a model which can help us rapidly estimate the distortion, called Standard Deviation for Distortion Estimation (STD-DE). Based on QZP-MDC, we focus on one of two received descriptions. Firstly, the distortion d can be represented as $(d_1 + d_2)/2$ where d_1 and d_2 denote the distortions of the two received descriptions. By the definition of mean squared error (MSE), d_1 and d_2 can be calculated by,

$$d_1 = \frac{1}{MN}\sum_{0\le x\le M}\sum_{0\le y\le N}(X(x,y)-X_1(x,y))^2$$
$$d_2 = \frac{1}{MN}\sum_{0\le x\le M}\sum_{0\le y\le N}(X(x,y)-X_2(x,y))^2 \quad (1)$$

where X_1 and X_2 are the reconstructions of the received descriptions and the original size of the frame is the width M and the height N. Next, d can be denoted as

$$d = \frac{1}{4MN}\sum_{0\le x\le M}\sum_{0\le y\le N}((2X(x,y)-X_+(x,y))^2 + X_-(x,y)^2) \quad (2)$$

where X_+ is the sum of $X_1 + X_2$ and X_- is the difference of $X_1 - X_2$. Then, X_+ can be computed as,

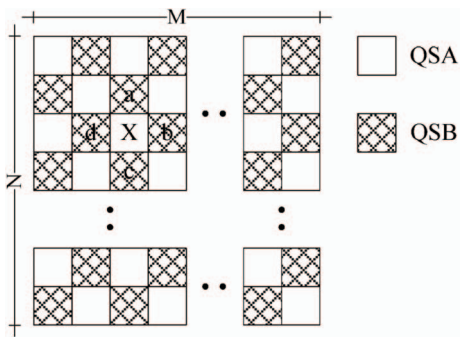


Figure 3. Quincunx pixel distribution and error concealment.

$$X_+ = \sum_{0\le u\le M}\sum_{0\le v\le N}(F_a(u,v)+F_b(u,v))a_{idct}(u,v)\alpha(u,x)\beta(v,y) \quad (3)$$

where F_a and F_b are the frequencies of two received descriptions, a_{idct} is parameter of inverse DCT, $\alpha(u, x) = \cos\frac{(2u+1)x\pi}{2M}$, and $\beta(v, y) = \cos\frac{(2v+1)y\pi}{2N}$. F_a and F_b can be deduced by Y_a and Y_b which are reconstructed with zero padding. Therefore, $F_a(u, v) + F_b(u, v) = \sum_{0\le u\le M_p}\sum_{0\le v\le N_p}(Y_a(x, y) + Y_b(x, y))\, a_{dct_p}(x, y)\alpha_p(x, u)\beta_p(y, v)$, where M_p is the width of a frame after padding zeros, N_p is the height of frame after padding zeros, a_{dct_p} is the parameter of DCT, $\alpha_p(u, x) = \cos\frac{(2u+1)x\pi}{2M_p}$, $\beta_p(v, y) = \cos\frac{(2v+1)y\pi}{2N_p}$. After that, we utilize nearby pixels to conceal errors next. Therefore,

$$Y_a = \begin{cases} Y(x,y), (x,y)\in QSA \\ \frac{1}{4}(Y(x-1,y)+Y(x+1,y)+Y(x,y-1)+Y(x,y+1)), (x,y)\in QSB \end{cases}$$
$$Y_b = \begin{cases} \frac{1}{4}(Y(x-1,y)+Y(x+1,y)+Y(x,y-1)+Y(x,y+1)), (x,y)\in QSA \\ Y(x,y), (x,y)\in QSB \end{cases} \quad (4)$$

where Y is the original frame with zero padding. Although there are some errors locating at the edge and corner of a frame, ignoring these errors can be generally accepted. Furthermore, $F_a(u, v) + F_b(u, v)$ can be deduced as

$$\begin{aligned} &F_a(u,v)+F_b(u,v) \\ &= \sum_{0\le u\le Mp}\sum_{0\le v\le Np}(Y(x,y)+\frac{1}{4}(Y(x-1,y)+Y(x+1,y) \\ &+Y(x,y-1)+Y(x,y+1)))a_{dct_p}(x,y)\alpha(x,u)\beta(y,v) \end{aligned} \quad (5)$$

where $Y(x, y) + \frac{1}{4}(Y(x-1, y) + Y(x+1, y) + Y(x, y-1) + Y(x, y+1))$ is represented as Y_+. Furthermore, $Y(x, y) = \sum_{0\le u\le Mp}\sum_{0\le v\le Np}F_p(u, v)\, a_{idctp}(u, v)\alpha_p(u, x)\beta_p(v, y)$, $Y_+ = \sum_{0\le u\le Mp}\sum_{0\le v\le Np}F_p(u, v)\, a_{idct_p}(u, v)\alpha_p(u, x)\beta_p(v, y)$ $(1 + \frac{1}{2}(\cos\frac{(2u+1)\pi}{2M_p} + \cos\frac{(2v+1)\pi}{2N_p}))$. Let $C_+ = 1 + \frac{1}{2}(\cos\frac{(2u+1)\pi}{2M_p} + \cos\frac{(2v+1)\pi}{2N_p})$, $F_a(u, v) + F_b(u, v) = F_p(u, v)\, C_+(u, v)$, so X_+ can be represented as,

$$X_+ = \sum_{0\le u\le M}\sum_{0\le v\le N}(F_p(u,v)C_+(u,v))a_{idct}(u,v)\alpha(u,x)\beta(v,y) \quad (6)$$

On the other hand,

$$X_- = \sum_{0\le u\le M}\sum_{0\le v\le N}(F_a(u,v)-F_b(u,v))a_{idct}(u,v)\alpha(u,x)\beta(v,y) \quad (7)$$

Based on the above equations, $F_a(u, v) - F_b(u, v)$ can be shown as

$F_a(u, v) - F_b(u, v) = \sum_{0\le u\le M_p}\sum_{0\le v\le N_p}(\gamma(x, y)(Y(x, y) - \frac{1}{4}(Y(x-1, y) + Y(x+1, y) + Y(x, y-1) + Y(x, y+1)))\, a_{dct_p}(x, y)\alpha(x, u)\beta(y, v)$. Let $C_- = 1 - \frac{1}{2}(\cos\frac{(2u+1)\pi}{2M_p} + \cos\frac{(2v+1)\pi}{2N_p})$, we can rewrite X_- as,

$$\begin{aligned} X_- = &\sum_{0\le u<M}\sum_{0\le v<N}((\sum_{0\le x<M_p}\sum_{0\le y<N_p}(r(x,y) \\ &\sum_{0\le u<M_p}\sum_{0\le y<N_p}(F_p(u,v)C_-(u,v)a_{idct_p}(u,v) \\ &\alpha_p(u,x)\beta_p(v,y))a_{dct_p}(x,y)\alpha_p(x,u)\beta_p(y,v))) \\ &a_{idct}(u,v)\alpha(u,x)\beta(v,y)) \end{aligned} \quad (8)$$

where $\gamma(x, y)$ be defined as $R(x, y) = \begin{cases} 1 & (x, y)\in QSA \\ -1 & (x, y)\in QSB \end{cases}$. Now, we have new equations to express X_+ and X_-. Finally, we can rewrite the distortion d as $d = EST_+ + EST_-$, where $EST_+ = \frac{1}{4MN}\sum_{0\le x\le M}\sum_{0\le y\le N}(2X(x, y) - X_+(x, y))^2$ and $EST_- = \frac{1}{4MN}\sum_{0\le x\le M}\sum_{0\le y\le N}X_-(x, y)^2$.

5. Triangular Linear Approximation Distortion Estimation

In previous section, we successfully derive STD-DE to estimate the distortion of a frame. However, the complexity of STD-DE is $O(n^2 \log n)$ which may cause too much time in computation. To address this issue, we pro-

pose Triangular Linear Approximation Distortion Estimation (TLA-DE) with the complexity of O(n) to expedite the computation at the little cost of the accuracy of distortion estimation. Figure 4(a) shows the MSEs of STD-DE, EST_+, EST_- in Foreman (CIF) with the width of the horizontally padded zeros between 16 and 688. To further elaborate the details of Figure 4(a), we discuss the MSEs of three critical ranges of the width of padded zeros, 16–80, 320–384, and 624–688, which are shown in Figures 4(b), 4(c), and 4(d), respectively. Figure 4(b) shows that the MSE of EST_- almost remains constant with respect to the increase of the width of padding zeros between 16 and 80. On the contrast, both the MSEs of EST_+ and STD-DE are gradually reduced along with the increase of the width of padded zeros. The MSE of STD-DE is the largest among these approaches. Figure 4(c) shows that when the number of zero padded equals the original width of the frame (352), the MSE of EST_- will be equal to that of EST_+, which is caused by $EST_+ = \frac{1}{4MN}\sum_{0\le x\le M}\sum_{0\le y\le N}(2X(x, y) - X_+(x, y))^2$ and $EST_- = \frac{1}{4MN}\sum_{0\le x\le M}\sum_{0\le y\le N}X_-(x, y)^2$. When the number of padded zeros equals the width of the original image, $2X - X_+$ will equal X_-. Figure 4(d) shows that the MSE of EST_- divided by that of EST_+ are reduced to nearly zero when the width of padded zeros is close to 688 which is almost two times of the width of the original frame. According to the above observations, EST_+ and EST_- would probably have the scale relation with respect to the width of padded zeros.

In order to further observe the relation, we compute the ratio of EST_- to EST_+ of 150 frames in 11 kinds of sequences, including Bridge-far (CIF), Bus (CIF), Container (CIF), Flower (CIF), Hall (CIF), Mobile (CIF), Mother-daughter (CIF), News (CIF), Silent (CIF), Tempete (CIF), and Waterfall (CIF). Figure 5 shows that when the number of zero padded equals the original width of frame, EST_- will be equal to EST_+. With the number of zero padding reducing or increasing, the ratio of EST_- to EST_+ will be almost linearly reduced and gradually close to zero.

Given that the ratio of EST_- to EST_+ is approximately linear, we can extimate EST_- with only known EST_+ to reduce the time complexity, which can be derived by a simple equation. Let the original width of a frame be σ

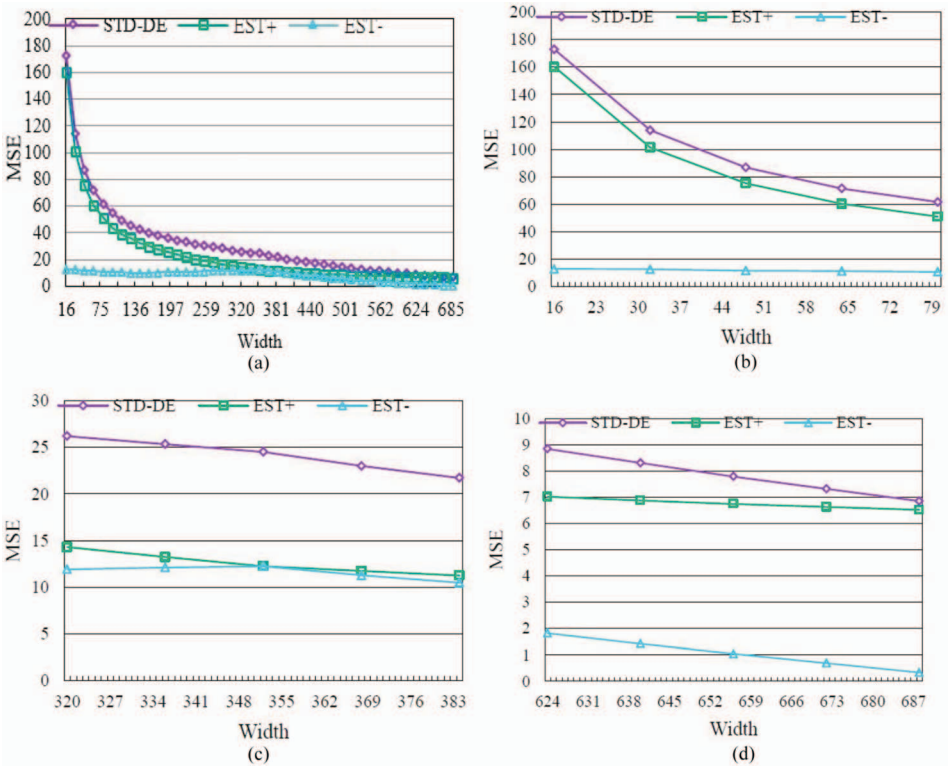


Figure 4. (a) Width of horizontally padded zeros between 16 to 688; (b) Width of horizontally padded zeros between 16 to 80; (c) Width of horizontally padded zeros between 320 to 384; (d) Width of horizontally padded zeros between 624 to 688.

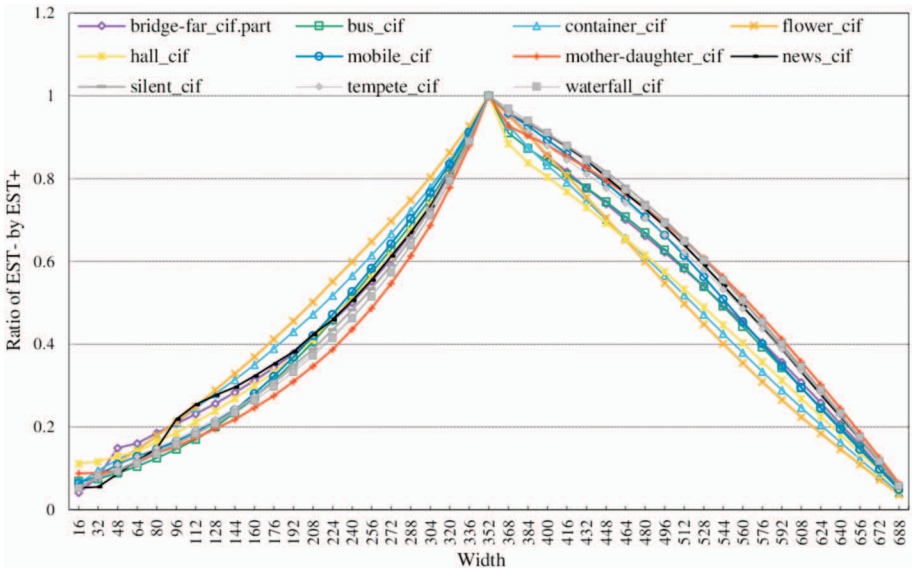


Figure 5. Ratio of EST_- by EST_+ of 150 frames in 11 kinds of sequences.

and its current width of zero padding be ρ. The ratio, $W = EST_+/EST_-$, can be defined as follows,

$$W = \begin{cases} 1+\dfrac{\rho}{\sigma} & \sigma < \rho \\ 1 & \sigma = \rho \\ 3-\dfrac{\rho}{\sigma} & \sigma > \rho \end{cases} \tag{9}$$

Given that the EST_+ for a frame with the number of padded zero equals 688, we can get W equal to 2.95 by the above equation. Then, the approximation of EST_- can quickly derived by EST_+ divided by W without using the actual equation by TLA-DE.

6. Finding the Acceptable Trade-off between Rate and Distortion

If the number of padded zeros increases, not only the average distortion of side information would be reduced but also the redundancy is increased. In order to search for the best number of padded zeros, we define Rate-Distortion Trade-off Indicator (RDTI) which contains Normalized Zero Padded (NZP) and Normalized Average Distortion of a description (NAD), where NZP represents the size of transferred data divided by the original data size and NAD represents that the average distortion of a description divided by the average distortion of a description with original size. The smaller the NZP and NAD, the better the efficiency of zero padding with lower distortion. In this paper, we assume that NZP and NAD are equally important. Therefore α is assigned to 0.5. The RDTI can be defined as follows,

$$RDTI = NZP \times \alpha + NAD \times (1-\alpha) \tag{10}$$

where α represents the weight. A smaller value of RDTI means that we can fewer number of padded zero with lower distortion for a description. Figure 6 RDTI of the first frame in Foreman (CIF) with the width of horizontally padding zeros for STD-DE.

The time complexity of computing RDTI for each number of padded zeros is proportional to the width of a frame. Searching for the best RDTI of a frame would be time consuming. In order to expedite the process, the searching can be stopped once an acceptable RDTI is obtained.

In this paper, we propose two methods to search for an acceptable RDTI, called Naïve Search and Best Divi-

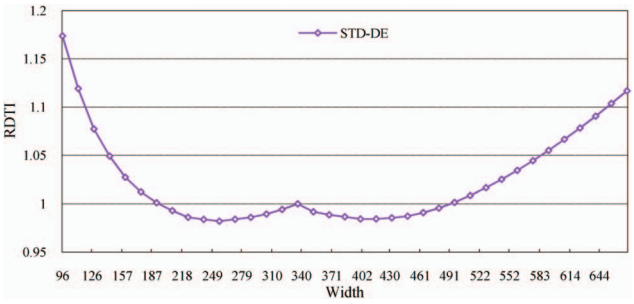


Figure 6. RDTI of the first frame in Foreman (CIF) with the width of horizontally padding zeros for STD-DE.

sion Search. Naïve search invokes one way naïve search twice to find two minimums of RDTI at two opposite directions starting from the width of a fame, and then select the minimum one. Naïve search algorithm and one way naïve search algorithm are shown in Figures 7 and 8. In addition, the best division search algorithm is similar to traditional binary search which is shown in Figure 9.

7. Experimental Results

In this section, we provide the analysis of the accuracy of STD-DE and TLA-DE, and also the efficiency of naïve search and the best division search based on STD-DE.

7.1 Errors of Distortion Estimation

We use errors of distortion estimation which is defined as $\left|\frac{a-b}{b}\right| \times 100\%$ to observe the accuracy of STD-DE, where a is the distortion by our estimation and b is the actual distortion. By definition, few errors of distortion estimation mean that the distortion estimation by STD-DE is close to the actual distortion. This experiment selected 150 frames from 11 kinds of sequences with the size of 352 × 288 as mentioned above. Figure 10 shows that the average error of distortion estimation for these sequences are less than 2%, meaning that STD-DE can accurately estimate distortion of zero padding. As shown in the figure, the largest gap of two points (352–368) in the figure is as small as 0.4% which can be treated as a neglectable anomaly and only 5 out of 11 clips experi-

Naïve search

Let L be the lower bound of padding width or height, U be the upper bound of padding width or height, Γ be the original padding width or height, and Δ be the step among each padding width or height.

1. Let R_ρ be the rate-distortion trade-off indicator with returning width or height ρ of one way naïve search with estimating padding width or height Γ upper bound of padding width or height U, and step among each padding width or height Δ.
2. Let R_σ be the rate-distortion trade-off indicator with returning width or height σ of one way naïve search with estimating padding width or height Γ upper bound of padding width or height L, and step among each padding width or height -Δ.
3. If $R_\rho \leqq R_\sigma$ then return ρ.
4. Else return σ.

Figure 7. Naïve search algorithm.

One way naïve search

Let ρ be the estimating padding width or height, P be the upper bound of padding width or height, and Δ be the step among each padding width or height.

1 Estimate the rate-distortion trade-off indicator with padding width or height ρ as R_ρ. if not estimated yet.
2 Estimate the rate-distortion trade-off indicator with padding width or height ρ + Δ as σ.
3 If $R_{\rho+\Delta} \geqq 1$ then return ρ.
4 Else if ρ + 2Δ equals to P then
  4.1 If $R_\rho \leqq R_{\rho+\Delta}$ then return ρ.
  4.2 Else return ρ + Δ.
5 Else
  5.1 Let R_σ be the rate-distortion trade-off indicator with returning width or height σ of one way naïve search with estimating padding width or height ρ + Δ, upper bound of padding width or height P, and step among each padding width or height Δ.
  5.2 If $R_\rho \leqq R_\sigma$ then return ρ.
  5.3 Else return σ.

Figure 8. One way naïve search algorithm.

Best division search

Let L be the lower bound of padding width or height, U be the upper bound of padding width or height.

1 If L equals to U than return the width or height L.
2 Let $m = \frac{1}{2} \times (L+U), \rho = \frac{1}{2} \times (L+m), \sigma == \frac{1}{2} \times (U+m)$.
3 Estimate the rate-distortion trade-off indicator with padding width or height ρ as R_ρ. if not estimated yet.
4 Estimate the rate-distortion trade-off indicator with padding width or height σ as R_σ. if not estimated yet.
5 If m equals to ρ or m equals to σ then
  5.1 If $R_\rho \leqq R_\sigma$ then return ρ.
  5.2 Else return σ.
6 If $R_\rho \leqq R_\sigma$ the return the returning width or height of best division search with lower bound of padding width or height L, upper bound of padding width or height m.
7 Else return the returning width or height of best division search with lower bound of padding width or height m, upper bound of padding width or height U.

Figure 9. Best division search algorithm.

ence such anomaly. Such a turning print on the error of STD-DE starts at the width of the original frame (352). Padding zeros whose width less than or equal to 352 means removing the high frequency of a sequence; on the other hand, padding zeros whose width larger than 352 means not only reserving all the frequencies of a frame but also increasing its spatial correlation. Obviously, the width of 352 is located the boundary of frequency variation whose impact on the distortion of a sequence will depend on its content, resulting in the inconsistence of the error of distortion estimation for STD-DE. Figure 11 shows that the worst error of distortion estimation by TLA-DE is up to 18%, but the average is 5.5% with 95% confidence level.

7.2 Efficiency of Search for Acceptable RDTI

Figure 12 shows the average time of searching for acceptable RDTI with naïve search and best division

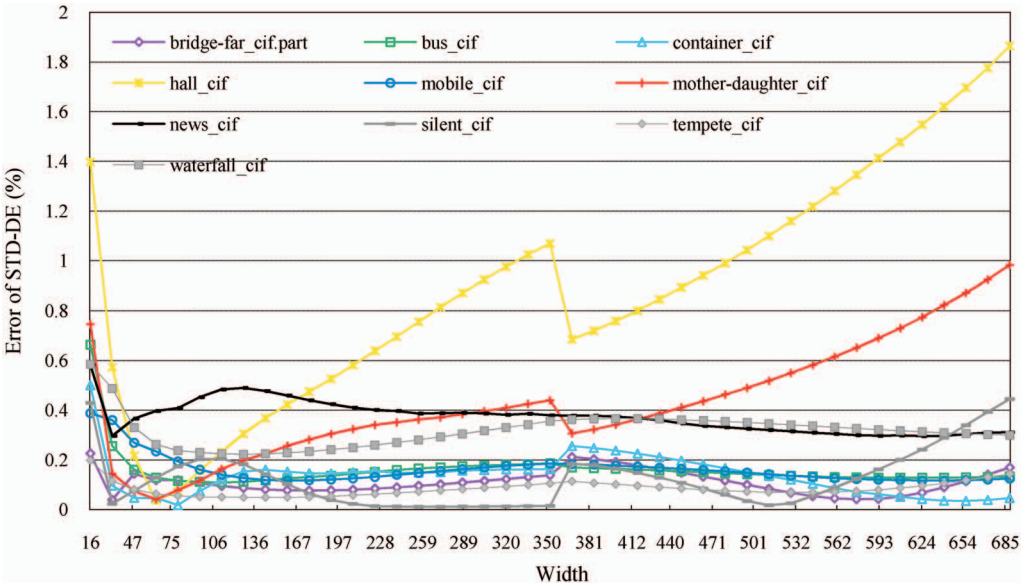


Figure 10. Average errors of distortion estimation of STD-DE in 11 kinds of sequences with horizontally padded zeros.

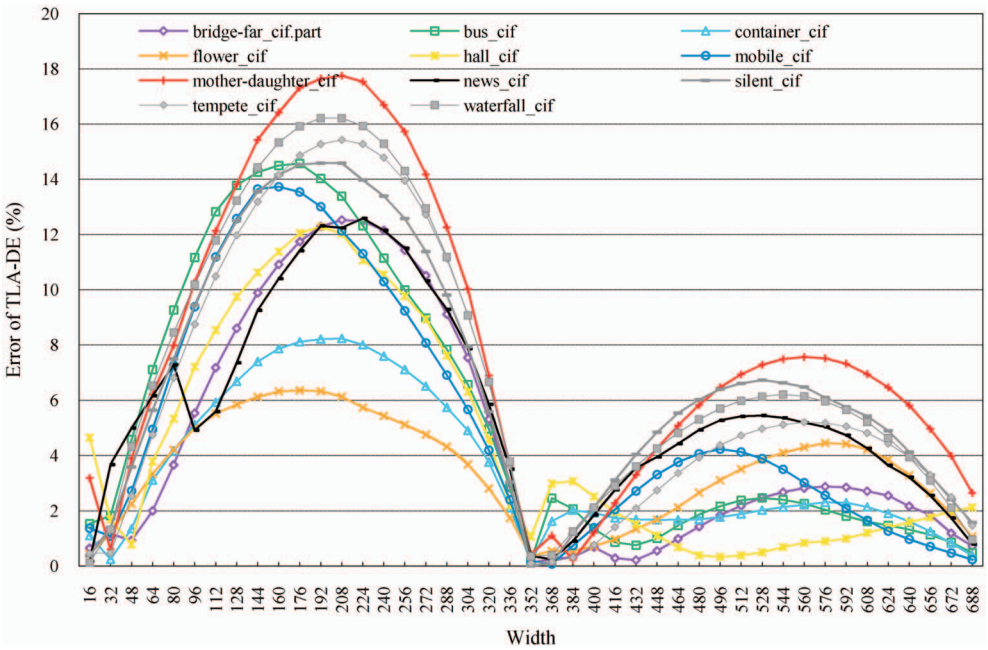


Figure 11. Average errors of distortion estimation of TLA-DE in 11 kinds of sequenceswith horizontally padded zeros.

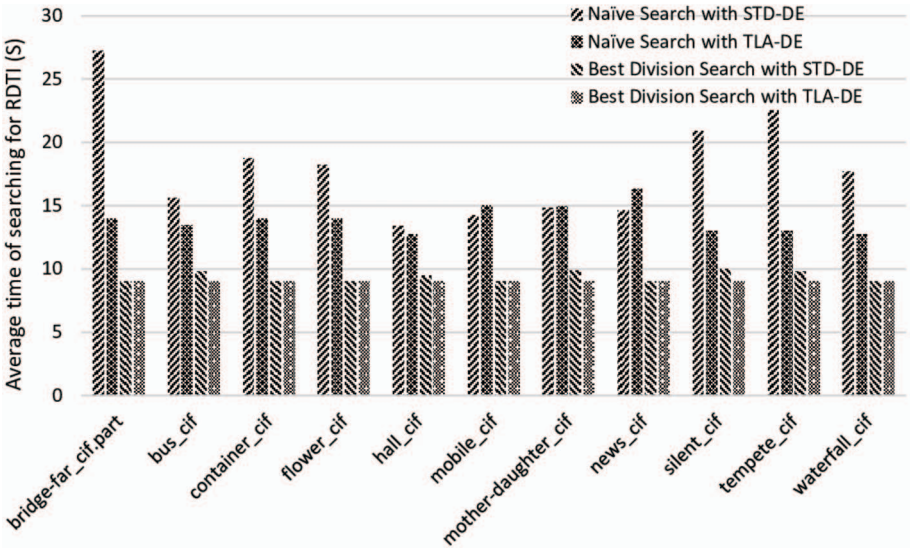


Figure 12. Average executing time of searching acceptable RDTI.

search under STD-DE and TLA-DE. In this experiment, we, also selected 150 frames of 11 kinds of above sequences. As shown in the figure, the average executing time by the best division search is about two times faster than that of naïve search either under STD-DE or TLA-DE. The best division search can help us effectively decrease the time for researching because it is based on binary search. On the contrary, the naïve search must search for the RDTI of each possible number of padded zeros, which may generate many unnecessary computations, resulting in poor searching efficiency.

8. Conclusions

In this paper, we propose two distortion estimations for MDC with zero padding based on spatial quincunx distribution, which are standard deviation for distortion estimation and triangular linear approximation distortion estimation (TLA-DE) to efficiently calculate the distortion for the number of padded zeros atthe horizontal direction. In order to search for the acceptable number of padded zeros, we propose two search methods, Naïve Search and Best Division Search, which are both based on the proposed rate-distortion trade-off indicator. Our experiments result show that the worst error of distortion estimation by TLA-DE is up to 18%, but the average is 5.5% with 95% confidence level. The executing time of best division search is about two times faster than that of naïve search for 11 test sequences. To sum up, we confirm that the acceptable trade-off between rate and distortion in low computational complexity can be found by using the best division search on TLA-DE. In the future, we plan to design a model to estimate the RDTI of the next frame based on the current RDTI in order to decide the number of padded zeros in the next frame.

Acknowledgements

This work was partially supported by National Science Council in Taiwan (R.O.C.) under grand MOST101-2221-E-024-023. The authors would like to thank the anonymous reviewers for their helpfuland constructive comments that greatly contributed to improving the final version of the paper. We also want to thank Ms. Tiffany Lin for her professionalproofreading.

References

[1] Bernardini, R., Durigon, M., Rinaldo, R., Celetto, L. and Vitali, A., "Polyphase Spatial Subsampling Multiple Description Coding of Video Streams with H264," *Proc. of International Conference on Image Processing*, Singapore, Vol. 5, pp. 3213–3216 (2004). doi: 10.1109/ICIP.2004.1421797

[2] Jia, J. and Kim, H. K., "Polyphase Downsampling Based Multiple Description Coding Applied to H.264 Video Coding," *Proc. of IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences*, Japanese, No. 6, pp. 1601–1606 (2006). doi: 10.1093/ietfec/e89-a.6.1601

[3] Wang, Y., Orchard, M. T., Vaishampayan, V. and Reibman, A. R., "Multiple Description Coding Using Pairwise Correlating Transforms," *IEEE Transactions on Image Processing*, Vol. 10, No. 3, pp. 351–366 (2001). doi: 10.1109/83.908500

[4] Goyal, V. K., "Multiple Description Coding: Compression Meets the Network," *IEEE Signal Processing Magazine*, Vol. 18, No. 5, pp. 74–93 (2002). doi: 10.1109/79.952806

[5] Wang, Y., Reibman, A. R. and Lin, S., "Multiple Description Coding for Video Delivery," *Proceedings of the IEEE*, Vol. 93, No. 1 pp. 57–70 (2005). doi: 10.1109/JPROC.2004.839618

[6] Kamnoonwatana, N., Agrafiotis, D. and Canagarajah, C. N., "Flexible Adaptive Multiple Description Coding for Video Transmission," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 22, No. 1, pp. 1–11 (2012). doi: 10.1109/TCSVT.2011.2129251

[7] Miao, Y., Yang, Q., Wu, C., Jiang, M. and Chen, J., "Multicast Virtual Network Mapping for Supporting Multiple Description Coding-based Video Applications," *Computer Networks*, Vol. 57, No. 4, pp. 990–1002 (2013). doi: 10.1016/j.comnet.2012.11.013

[8] Hsiao, C. W. and Tsai, W. J., "Hybrid Multiple Description Coding Based on H.264," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 20, No. 1, pp. 76–87 (2010). doi: 10.1109/TCSVT.2009.2026973

[9] Tillo, T., Grangetto, M. and Olmo, G., "Redundant Slice Optimal Allocation for H.264 Multiple Description Coding," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 18, No. 1, pp. 59–70 (2008). doi: 10.1109/TCSVT.2007.913751

[10] Radulovic, I., Frossard, P., Wang, Y. K. and Hannuksela, M. M., "Multiple Description Video Coding With H.264/AVC Redundant Pictures," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 20, No. 1, pp. 144–148 (2010). doi: 10.1109/TCSVT.2009.2026815

[11] Choupani, R., Wong, S. and Tolum, M., "Spatial Multiple Description Coding for Scalable Video Streams," *International Journal of Digital Multimedia Broadcasting*, Vol. 2014, No. 132621 (2014). doi: 10.1155/2014/132621

[12] Radulovic, I., Wang, Y. K., Wenger, S., Hallapuro, A., Hannuksela, M. M. and Frossard, P., "Multiple Description H.264 Video Coding with Redundant Pictures," *Proc. of the International Workshop on Workshop on Mobile Video*, New York, U.S.A., pp. 37–42 (2007).

[13] Reibman, A. R., Jafarkhani, H., Wang, Y., Orchard, M. T. and Rohit, P., "Multiple-Description Video Coding Using Motion-Compensated Temporal Prediction," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 12, No. 3, pp. 193–204 (2002). doi: 10.1109/76.993440

[14] Wang, D., Canagarajah, N., Redmill, D. and Bull, D., "Multiple Description Video Coding Based on Zero Padding," *Proc. of the International Symposium on Circuits and Systems*, Vancouver, Canada, pp. II - 205-8 (2004). doi: 10.1109/ISCAS.2004.1329244

[15] Wang, D., Canagarajah, N., Agrafiotis, D. and Bull, D., "Error Concealment for Slice Group based Multiple Description Video Coding," *Proc. of IEEE International Conference on Image Processing*, Vol. 1, pp. I-769-72 (2005). doi: 10.1109/ICIP.2005.1529864

[16] Farzamnia, A., Yusof, S. K. S., Nathaniel, S. and Lee, R. M. A., "Recovering Single Lost Description Using DCT-MDC Based on Network Coding," *Proc. of International Conference onTelematics and Future Generation Networks*, Kuala Lumpur, pp. 59–64 (2015). doi: 10.1109/TAFGEN.2015.7289576

[17] Kamnoonwatana, N., "Format-compliant Perceptual Video Encryption Based on Multiple Description Coding," *Proc. of IEEE Malaysia International Conference on Communications*, Kuala Lumpur, pp. 75–80 (2013). doi: 10.1109/MICC.2013.6805803

[18] Tsai, W. J. and Chen, J. Y., "Joint Temporal and Spatial Error Concealment for Multiple Description Video Coding," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 20, No. 12, pp. 1822–1833 (2010). doi: 10.1109/TCSVT.2010.2087816

[19] Ma, M., Au, O. C., Chan, S. H. G. and Sun, M. T., "Edge-Directed Error Concealment," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 20, No. 3, pp. 382–395 (2010). doi: 10.1109/TCSVT.2009.2035839

[20] Wang, A. H., Zhao, Y. and Bai, H. H., "Robust Multiple Description Distributed Video Coding Using Optimized Zero-padding," *Science in China Series F: Information Sciences*, Vol. 52, No. 2, pp. 206–214 (2009). doi: 10.1007/s11432-009-0037-5

[21] Shirani, S., Gallant, M. and Kossentini, F., "Multiple

Description Image Coding Using Pre- and Post-processing," *Proc. of International Conference on Information Technology: Coding and Computing*, Las Vegas, NV, pp. 35–39 (2001). doi: 10.1109/ITCC.2001.918761

[22] Olmo, G. and TILLO, T., "Directional Multiple Description Scheme for Still Images," *Proc. of 2003 10th IEEE International Conference on Electronics, Circuits and Systems*, Sharjah, United Arab Emirates, pp. 886–889 (2003). doi: 10.1109/ICECS.2003.1301929

Manuscript Received: Sep. 18, 2015
Accepted: Mar. 21, 2016