

IDEMPOTENT H.264 INTRAFRAME MULTI-GENERATION CODING

Ziyuan Zhu, Student Member, IEEE and Tao Lin, Member, IEEE

VLSI Lab., Tongji University, Shanghai 200092, China

ABSTRACT

This paper studies the idempotence of H.264 intraframe multi-generation coding. We analyze the H.264 transform and quantization and reveal that for some quantization parameters there always exists at least one clipping compensation matrix to make the H.264 intraframe multi-generation coding idempotent as long as the same prediction mode and the same quantization parameter are selected for each coding generation. In addition, an idempotent H.264 intraframe multi-generation coding procedure is presented.

Index Terms—H.264, multi-generation, idempotent

1. INTRODUCTION

Multi-generation coding is a repeated compression and decompression process of videos [1]. In the case of lossless compression, the re-compressed and re-decompressed video sequence is exactly identical to the original one. For lossy compression, it can be idempotent, i.e. the video sequence reconstructed at each decompression cycle is identical to the one after the first decompression stage. In some applications, e.g. a video editing system, idempotence is desirable. However the idempotence is usually not satisfied for the current JPEG and MPEG standards [2], [3]. Nowadays, the H.264 applications range from portable devices to HDTV. As the office/home multimedia networks become pervasive, it becomes more common that video contents need to be encoded and decoded multiple times during their lifetime [4], [5]. Therefore, idempotent H.264 coding is becoming an important issue. The idempotence of H.264 like hybrid video coding has been studied in [5] and it is shown that H.264 multi-generation coding can have 2 ~ 4 dB PSNR drop after 3 generations in some cases. This paper analyzes the H.264 quantization (Qu), transform (Tr), inverse-transform (ITr) and dequantization (DQu), and proves the quadruple (Qu, Tr, ITr, DQu) to be idempotent for some quantization parameter (QP). Based on the idempotence of the quadruple (Qu, Tr, ITr, DQu), we present an idempotent H.264 coding architecture. In addition, an idempotent H.264 intraframe multi-generation coding procedure is presented.

Section 2 gives the fundamental of idempotent H.264 coding. In section 3, we present an idempotent H.264 intraframe coding procedure, followed by experimental results in section 4. Conclusions are given in section 5.

2. FUNDAMENTAL OF IDEMPOTENT H.264

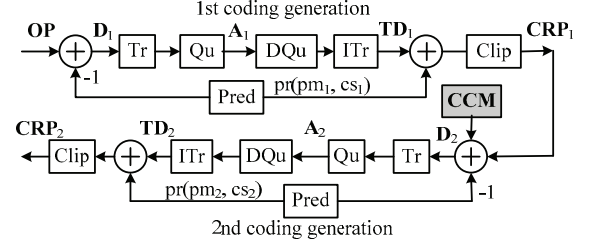


Fig. 1. Idempotent H.264 multi-generation architecture

Without loss of generality, we consider only the 1st and 2nd coding generations of the idempotent H.264 intraframe multi-generation coding architecture proposed in Fig. 1. In this paper, we assume the deblocking loop filter is disabled and discuss baseline and main profile cases. Since entropy encoding and decoding are lossless, we have removed them for simplicity. **OP** is the original picture. **CRP₁** and **CRP₂** are reconstructed pictures of the 1st and 2nd coding, respectively. $pr(pm, cs)$ is the intra prediction result depending on the prediction mode (pm) and the previously reconstructed neighboring samples (cs). Tr , ITr , Qu and DQu are defined as follows [6]:

$$Tr(\mathbf{X}) = \mathbf{H}\mathbf{X}\mathbf{H}^T \quad (1)$$

$$ITr(\mathbf{X}) = (\mathbf{H}_{inv}\mathbf{X}\mathbf{H}_{inv}^T) // 64 \quad (2)$$

$$Qu(\mathbf{X}, QP) = \text{sign}(\mathbf{X}) (|\mathbf{X}| \otimes \mathbf{M}(QP) + f) // 2^{15+qbits} \quad (3)$$

$$DQu(\mathbf{X}, QP) = 2^{qbits} \otimes \mathbf{X} \otimes \mathbf{V}(QP) \quad (4)$$

where \otimes denotes element-by-element multiplication, $\mathbf{M}(QP)$ and $\mathbf{V}(QP)$ are quantization and de-quantization matrices indexed by QP , $qbits = QP // 6$, f is equal to $2^{15+qbits} // 3$ or equal to $2^{16+qbits} // 3$ for the DC quantization, \mathbf{H} and \mathbf{H}_{inv} are forward and pseudo-inverse transform matrices [6], the symbol $//$ denotes division with rounding towards minus infinity and $///$ denotes division with rounding towards nearest integer. It can be seen that the major difference between the proposed idempotent H.264 coding and the traditional H.264 coding is that we introduce a clipping compensation matrix (**CCM**) during the 2nd encoding generation. The value of the **CCM** will be found during the 2nd encoding generation to make the coding idempotent, i.e.:

$$\mathbf{CRP}_1 = \mathbf{CRP}_2. \quad (5)$$

All the signals, such as **OP**, **CRP₁**, **CRP₂** and **CCM** are typically in the format of matrices with finite word length.

2.1. The existence of CCM

There always exists a **CCM** to make the H.264 multi-generation coding system shown in Fig. 1 idempotent, if during the 2nd encoding, QP_2 and pm_2 are selected to satisfy:

$$QP_2 = QP_1, pm_2 = pm_1 \quad (6)$$

and the quadruple (Qu, Tr, ITr, DQu) of the coding system satisfies the idempotence:

$$Qu(Tr(ITr(DQu(A_1, QP_1))), QP_2) = A_1 \quad (7)$$

where QP_1 and pm_1 are the quantization parameter and prediction mode of the 1st coding generation, and QP_2 and pm_2 are those of the 2nd coding generation.

First, notice that $cs_1 = cs_2$, if the previously reconstructed neighboring macroblocks are idempotent. So, in the case of $pm_2 = pm_1$, we have

$$pr(pm_1, cs_1) = pr(pm_2, cs_2). \quad (8)$$

Applying ITr, DQu to both sides of (7) and hiding QP_1 and QP_2 which are equal to each other, we get

$$ITr(DQu(Qu(Tr(ITr(DQu(A_1)))))) = ITr(DQu(A_1)). \quad (9)$$

Substituting $TD_1 = ITr(DQu(A_1))$ into (9) yields

$$ITr(DQu(Qu(Tr(TD_1)))) = TD_1. \quad (10)$$

From Fig. 1, we have

$$D_2 = CCM + CRP_1 - pr(pm_2, cs_2) \quad (11)$$

$$CRP_1 = clip(TD_1 + pr(pm_1, cs_1)) \quad (12)$$

$$CRP_2 = clip(ITr(DQu(Qu(Tr(D_2)))) + pr(pm_2, cs_2)). \quad (13)$$

Substituting (10) into (12) gives

$$CRP_1 = clip(ITr(DQu(Qu(Tr(TD_1)))) + pr(pm_1, cs_1)). \quad (14)$$

Comparing (13) with (14), and considering (8), we only need $D_2 = TD_1$ to satisfy $CRP_1 = CRP_2$. Then from (11), we obtain **CCM** given by $CCM = TD_1 + pr(pm_1, cs_1) - CRP_1$.

Therefore, there always exists at least one **CCM** to make the multi-generation coding system idempotent, if (6) and (7) are satisfied.

2.2 The idempotence of the quadruple

Eq. (7) reveals a basic and important property that the quadruple must have for idempotence. It can be shown that for certain QP values, (7) does hold.

Generally, a finite precision division can be factored into an infinite precision division and a rounding fraction:

$$c = a/b = a/b - frac_{div} \quad 0 \leq frac_{div} < 1 \quad (15)$$

$$c = a/b = a/b + frac_{rou} \quad -0.5 \leq frac_{rou} \leq 0.5 \quad (16)$$

Luma 4×4 Case

From (2), (3), (15) and (16), and without loss of generality assuming $X \geq 0$, (3) and (2) become

$$Qu(X, QP) = (X \otimes M(QP) + f)/2^{15+qbits} - frac_q \quad (17)$$

$$ITr(X) = ((H_{inv} \bullet X + frac_1) \bullet H_{inv}^T + frac_2) / 64 + frac_3 \quad (18)$$

where \bullet denotes the matrix multiplication with infinite precision. $0 \leq frac_q < 1$, $-0.5 \leq frac_1, frac_2, frac_3 \leq 0.5$. We can rewrite (18) as

$$ITr(X) = H_{inv} \bullet X \bullet H_{inv}^T / 64 + frac_{inv} \quad (19)$$

where $|frac_{inv}| = |(frac_1 \bullet H_{inv}^T + frac_2) / 64 + frac_3| \leq 137/256$.

Substituting (19) into (1), and simplifying,

$$Tr(ITr(X)) = X \otimes K / 64 + frac_t \quad (20)$$

where $K = \begin{bmatrix} 16 & 20 & 16 & 20 \\ 20 & 25 & 20 & 25 \\ 16 & 20 & 16 & 20 \\ 20 & 25 & 20 & 25 \end{bmatrix}$, and

$$|frac_t| = |H \bullet (frac_{inv}) \bullet H^T| \leq \begin{bmatrix} 16 & 24 & 16 & 24 \\ 24 & 36 & 24 & 36 \\ 16 & 24 & 16 & 24 \\ 24 & 36 & 24 & 36 \end{bmatrix} \otimes frac_{inv}.$$

Substituting (4) and (20) into (17) and hiding QP_1 and QP_2 which are equal to each other, we get

$$Qu(Tr(ITr(DQu(A_1)))) = A_1 + FRAC \quad (21)$$

where

$$FRAC = A_1 \otimes V(QP) \otimes K \otimes M(QP) / 2^{21} - A_1 + frac_t \otimes M(QP) / 2^{15+qbits} + f / 2^{15+qbits} - frac_q.$$

In (21), since all elements of $Qu(Tr(ITr(DQu(A_1))))$ and A_1 are integers, it only needs $-1 < FRAC < 1$ to satisfy (7). Assume the coding bitdepth is 8-bit, then the residual $|D_1| \leq 255$. From Fig. 1, we have

$$|A_1| \leq Qu\left(\begin{bmatrix} 255 & 255 & 255 & 255 \\ 255 & 255 & 255 & 255 \\ 255 & 255 & 255 & 255 \\ 255 & 255 & 255 & 255 \end{bmatrix} \otimes \begin{bmatrix} 16 & 24 & 16 & 24 \\ 24 & 36 & 24 & 36 \\ 16 & 24 & 16 & 24 \\ 24 & 36 & 24 & 36 \end{bmatrix}, QP\right).$$

Since M , V and K are all constant matrices. Substituting all the QP values into **FRAC**, it can be verified that (7) is satisfied when $QP_1 = QP_2 \geq 21$.

Chroma 8×8 Case

In the case of chroma 8×8, the AC coefficients can be verified as same as those of luma 4×4 ones. However, the case of 2×2 DC coefficients is different.

For the quantization of the DC coefficients matrix $A_{DC2 \times 2}$, firstly the 2×2 Hadamard transform is applied:

$$Z_{2 \times 2} = G_{2 \times 2} A_{DC2 \times 2} G_{2 \times 2}^T \quad (22)$$

where $G_{2 \times 2}$ is 2×2 Hadamard matrix. Then the 2×2 DC quantization is applied; that is,

$$Qu_{DC}(Z_{2 \times 2}) = sign(Z_{2 \times 2}) (|Z_{2 \times 2}| \otimes M_{DC}(QP) + f) / 2^{16+qbits}. \quad (23)$$

For the dequantization of $A_{DC2 \times 2}$, firstly, the 2×2 Hadamard transform defined in (22) is applied. Then the 2×2 DC dequantization is applied; that is,

$$DQu_{DC}(Z_{2 \times 2}) = \begin{cases} 2^{qbits-1} \otimes Z_{2 \times 2} \otimes V_{DC}(QP) & QP \geq 6 \\ Z_{2 \times 2} \otimes V_{DC}(QP) / 2 & QP < 6 \end{cases}. \quad (24)$$

From (20), we get the pair (Tr, ITr); that is

$$Tr(ITr(Z_{2 \times 2})) = Z_{2 \times 2} \otimes K_{DC} / 64 + frac_t \quad (25)$$

where $K_{DC} = \begin{bmatrix} 16 & 16 \\ 16 & 16 \end{bmatrix}$, M_{DC} and V_{DC} are DC quantization

and de-quantization matrices indexed by QP. From (22) to (25) and (15), (16), it can be verified that for all $QP_1 = QP_2 \geq 27$, $Qu_{DC}(Tr(ITr(DQu_{DC}(A_{DC2 \times 2})))) = A_{DC2 \times 2}$.

So, in conclusion, for chroma 8×8, (7) is satisfied when $QP_1 = QP_2 \geq 27$.

Luma 16×16 Case

Similar to the case of chroma 8×8, it can be verified that for luma 16×16, (7) is satisfied when $QP_1 = QP_2 \geq 33$.

Therefore, for the proposed H.264 coding architecture, if $QP \geq 21$, $QP \geq 27$ and $QP \geq 33$ for luma 4×4, chroma 8×8, and luma 16×16, respectively, then there always exists a **CCM** to achieve idempotence as long as the 2nd coding has the same QP and prediction mode as the 1st coding. For example, a two generation coding of a 4×4 block in the QCIF sequence Mobile is shown in Fig. 2, $QP_1 = QP_2 = 21$ and $pm_1 = pm_2 = \text{Horizontal}$. (a) shows the original block, **OP** in Fig. 1, and the reconstructed neighboring samples. (b) shows the 1st coding residual D_1 . (c) is the reconstructed block before clipping, $TD_1 + pr(pm_1, cs_1)$. (d) is the decoded block of the 1st coding CRP_1 which is also the input of the 2nd coding. (e) is the 2nd coding residual. There exists a **CCM** = $TD_1 + pr(pm_1, cs_1) - CRP_1$ shown in (f). (g) is the clipping compensated residual D_2 given by $D_2 = CCM + (e)$. (h) shows the decoded result of the 2nd coding CRP_2 by encoding and decoding D_2 . It can be seen that the two generation coding is idempotent since (h) is identical to (d).

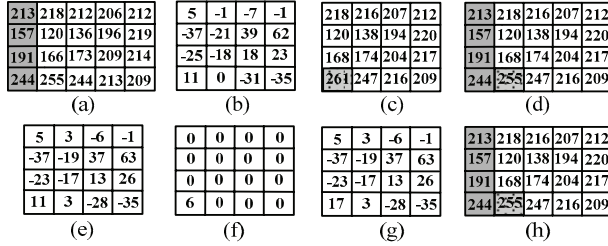


Fig. 2. Illustration of a two generation coding

3. AN IDEMPOTENT H.264 CODING PROCEDURE

Based on the discussion of section 2, this section presents an idempotent H.264 intraframe coding procedure as shown in Fig. 3. The procedure has three additional steps to achieve idempotence: Prediction Mode and QP Restriction (PMQR), Prediction Mode and QP Identification (PMQI) and Clipping Compensation (CC).

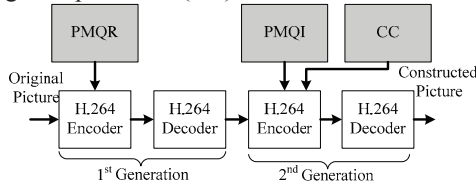


Fig. 3. Idempotent H.264 coding procedure

3.1. PMQR

In general, a macroblock (MB) can be encoded using either luma 16×16 or luma 4×4 together with chroma 8×8 and the QP can take values from 0 to 51. However, for idempotence, PMQR must be applied during the 1st encoding. PMQR means that the luma $QP \geq 21$, the chroma $QP \geq 27$ and only Intra 4×4 prediction mode can be used when $21 \leq QP < 33$.

3.2 PMQI and CC

PMQI and CC are used to find pm_2 , QP_2 and **CCM** to satisfy (5). From (11) and Fig. 1, (5) is equivalent to $\text{clip}(TD_2 + \text{pred}(pm_2, cs_2)) = D_2 - \text{CCM} + \text{pred}(pm_2, cs_2)$. (26)

Considering the clipping operation, (26) is equivalent to

$$\begin{cases} TD_2[i, j] \geq D_2[i, j] - CCM[i, j], & CRP_1[i, j] = 255 \\ TD_2[i, j] \leq D_2[i, j] - CCM[i, j], & CRP_1[i, j] = 0 \\ TD_2[i, j] = D_2[i, j] - CCM[i, j], & \text{Others} \end{cases}$$

We define a function $SD_{n \times n}(pm, QP, CCM)$ as follows:

$$SD_{n \times n}(pm_2, QP_2, CCM) = \sum_{i,j=0}^n SEEC[i, j] \quad (27)$$

where $n = 4, 16$ and 8 for luma 4×4, luma 16×16 and chroma 8×8 respectively, and

$$SEEC[i, j] = \begin{cases} u(TD_2[i, j] - D_2[i, j] + CCM[i, j]) & CRP_1[i, j] = 255 \\ u(D_2[i, j] - TD_2[i, j] - CCM[i, j]) & CRP_1[i, j] = 0 \\ (D_2[i, j] - TD_2[i, j] - CCM[i, j])^2 & \text{Others} \end{cases}$$

$$\text{where } u(x) = \begin{cases} 0 & x \geq 0 \\ -x & \text{Others} \end{cases}$$

Since (5) is equivalent to $SD_{n \times n} = 0$, PMQI and CC are to find a triple of pm_2 , QP_2 and **CCM** to satisfy $SD_{n \times n} = 0$ instead of (5). If no clipping occurs, i.e. **CCM** = 0, one obvious method to identify pm_2 and QP_2 is to try all possible pm_2 and QP_2 , which make $SD_{n \times n} = 0$. Then search through all possible prediction modes and QP to evaluate a traditional mode selection criterion, e.g. SAD, and select one candidate which minimize the SAD. However, clipping always occurs in practice. Thus, one key issue of idempotent coding is to find the **CCM**. One way is to find the root of equation $SD_{n \times n} = 0$. But it's hard to solve such a high dimensional nonlinear equation. Experimental results show that there are more than one **CCM** satisfying $SD_{n \times n} = 0$. So, it's easier to formulate the solution of **CCM** as an optimization problem. In this paper, the genetic algorithm (GA) [7] is employed. The flow chart of PMQI and CC is shown in Fig.4. In this implementation, the PMQI and CC based on GA are only applied in the unit of 4×4 sub-blocks for efficiency. The cost function [7] is (27). The chromosome [7] includes pm_2 and **CCM**. First, the MB is partitioned into 4×4 sub-blocks. If all the sub-blocks clipped, the current MB will be encoded as I_PCM mode for the GA optimization efficiency. Otherwise, PMQI finds the QP of the current MB by searching through all the possible prediction mode and QP in the no clipped sub-blocks. Then GA-based PMQI and CC are applied for the clipped sub-blocks. It turn out that even some clipped sub-blocks have $SD_{n \times n} = 0$ when **CCM** = 0, $pm_2 = pm_1$ and $QP_2 = QP_1$, thus GA is not needed. We name these sub-blocks GA bypass ones if a pm_2 satisfying $SD_{n \times n} = 0$ can be found by searching through all the possible prediction modes when **CCM** = 0.

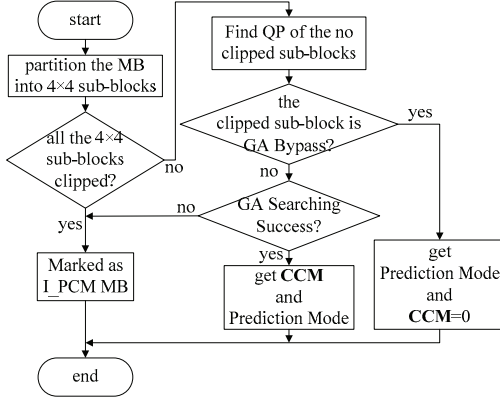


Fig. 4. PMQI and CC flow chart

4. EXPERIMENTAL RESULTS

We modified JVT JM14 to develop an idempotent H.264 encoder. The new encoder performs PMQR in the 1st coding generation and PMQI and CC instead of the traditional prediction mode selection in the higher coding generations. Using the idempotent encoder, two 720P test sequences Crew and Parkrun are compressed 10 generations consecutively in the following two experiments comparing with JM14 in the same multi-generation environments. The performances of GA-based PMQI and CC are evaluated.

Experiment A

Coding parameters: ChromaQPOffset = 6, Bitrate = 9.5 Mbps, CAVLC, RDO = 1, Error metric for Mode Decision = SAD, Disable Deblocking Filter, IntraPeriod = 1, FrameRate = 30. The test sequence is the first 50 frames of Crew. In this experiment, Luma QP = 21, 23, 24, 25, 26, 27, 31 and 35. Chroma QP = 27, 29, 30, 31, 32, 34 and 36. The PMQR ensures only the Intra 4x4 prediction modes are selected when QP < 33. The asterisk line of Fig. 5 illustrates PSNR performance of the proposed idempotent coding procedure while the square line shows that of JM14. Fig. 5 shows that the idempotent H.264 coding procedure has no quality degradation when doing re-compression while JM14 brings about 2.69 dB PSNR degradation after 9 generations. It also shows that PMQR cause very small PSNR difference (0.02 dB) between the idempotent coding and the 1st coding of JM14. There is no clipping occurring in the experiment.

Experiment B

Coding parameters: ChromaQPOffset = 6, Bitrate = 35 Mbps, CAVLC, RDO = 1, Error metric for Mode Decision = SAD, Disable Deblocking Filter, IntraPeriod = 1, FrameRate = 30. The test sequence is the first 50 frames of Parkrun. In this experiment, Luma QP = 25, 29, 32, 33, 34 and 36. Chroma QP = 30, 33, 35, 36 and 37. Fig. 6 shows that JM14 has about 3.62 dB PSNR degradation after 9 generations while the idempotent encoder has no quality degradation. Table I shows that there are 7951 clipped sub-blocks, and 7947 of them are GA bypass ones. There are 4 sub-blocks need GA searching to find a CCM. All the GA

optimization routines successfully return within 50 GA generations.

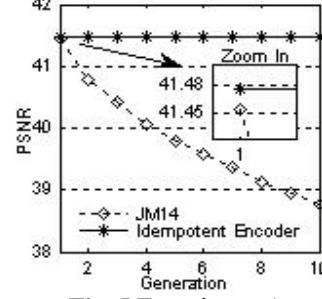


Fig. 5. Experiment A

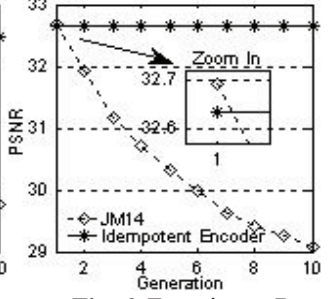


Fig. 6. Experiment B

Table I PERFORMANCE OF CC IN EXPERIMENT B

Clipped Sub-block(s)	GA Bypassed	GA Success	I_PCM MB
7951	7947	4	0

5. CONCLUSION

In this paper, we analyze the idempotence of H.264 and present an idempotent H.264 intraframe multi-generation coding architecture. Based on the proposed architecture, we present an idempotent H.264 coding procedure. Experiment results show that for some of the coding parameters, the proposed idempotent H.264 coding procedure produces no generation PSNR loss. The idempotent H.264 encoder has some restriction on QP values, but most QP values used in common applications are still allowed and the impact of the restriction on coding efficiency is minor. Our further work will focus on the idempotent multi-generation coding both in intra and inter frames with deblocking loop filter.

6. REFERENCES

- [1] C. Horne, T. Naveen, and A. Tabatabai, "Study of the characteristics of the MPEG2 4:2:2 profile-application of MPEG2 in studio environment", *IEEE Trans. Circuits and Systems for Video Technology*, Vol. 6, no. 3, pp. 251-272, 1996.
- [2] A.T. Erdem, and M.I. Sezan, "Multi-generation characteristics of the MPEG video compression standards", *IEEE Int. Conf. Image Processing*, vol 2, pp. 933-937, 1994.
- [3] R.L. Joshi, M. Rabbani, and M.A. Lepley, "Comparison of multiple compression cycle performance for JPEG and JPEG2000", *Proc.SPIE*, vol. 4115, pp. 492-501, 2000
- [4] R. Krishnamoorthy, "High definition, anywhere: How ultra wideband makes wireless HDMI possible", *2007 4th Annual IEEE Consumer Communications and Networking Conf.*, pp. 395-399
- [5] T. Lin, "Achieving Re-Loss-Free Compression in H.264 Like Hybrid Video Coding", *IEEE, Int. Conf. Signals, Circuits & Systems*, Hammamet, Tunisia, Nov. 2008, to be published
- [6] H. S. Malvar, A. Hallapuro, M. Karczewicz, and L. Kerofsky, "Low-complexity transform and quantization in H.264/AVC", *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 598-603, Jul. 2003.
- [7] R.L. Haupt and S.E. Haupt, *Practical Genetic Algorithms*, 2nd ed., John Wiley & Sons, New York, 2004.