A COMPUTING METHOD FOR LINEAR CONVOLUTION IN THE DCT DOMAIN

Izumi ITO and Hitoshi KIYA†

Tokyo Institute of Technology
2-12-1, Ookayama, Meguro-ku, 152-8552, Tokyo, Japan
phone: + 81-3-5734-3232, fax: +81-3-5734-2560,
email: ito@nh.cradle.titech.ac.jp

†Tokyo Metropolitan University
6-6, Asahigaoka, Hino-shi, 191-0065, Tokyo, Japan
phone: + 81-42-585-8419, fax: +81-42-583-5119,
email: kiya@sd.tmu.ac.jp

ABSTRACT

We propose a computing method for linear convolution between sequences using discrete cosine transform (DCT). Zero padding is considered as well as linear convolution using discrete Fourier transform (DFT). Analyzing the output range of the resulting convolution, we derive the minimum number of zero-padding before and after the sequences. The proposed method requires DCT-2 and DCT-1 transforms regardless of sequences, and can calculate linear convolution with both linear phase filter and non-linear phase filter. The computational complexity of the proposed method is lower than that of linear convolution using DFT. In addition, the proposed method can be used for computation of linear correlation between two sequences.

1. INTRODUCTION

Discrete cosine transform (DCT) is closely related to discrete Fourier transform (DFT) [1]. One of the advantages of DCT over DFT is calculation with real numbers. Generally, DFT requires calculation with complex numbers. Since the operations of complex numbers take more than those of real numbers, the number of calculation of DCT is less than that of DFT for the same length of sequences. DCT has sets of basis of cosines, which are periodic and even symmetry [2]. Involving several ways of symmetry, DCT has some types. DCT type 2 (DCT-2) is most popular because it is used for compression standard, such as JPEG and MPEG. Accordingly many implementations of DCT-2 have been developed.

There are some computing methods for convolution using DCT [3]-[5]. 40 types of symmetric convolution and their convolution multiplication properties in the DCT and discrete sine transform (DST) domains are derived [3]. The type of transforms to be used depends on the type of the symmetry of the sequences to be convolved in symmetric convolution. Symmetric convolution cannot calculate linear convolution with non-linear phase filter. For linear convolution with non-linear phase filter using DCT, Reju, Koh, and Soon proposed a method using DCT-2 and DST-2 as forward transform for a sequence [4], and Suresh and Sreenivas proposed a method using DCT-4 and DST-4 [5]. These methods require DST in addition to DCT as both forward and inverse transforms.

In the present paper, we propose a computing method for linear convolution using DCT-2 as forward transform for a sequence and DCT-1 as inverse transform regardless of sequences. The proposed method can calculate linear convolution with both linear phase filter and non-linear phase filter. As well as linear convolution using DFT, we consider proper zero-padding in the spatial domain. Due to the symmetry implied by DCT, the resulting consists of four convolutions. We analyze the each output range of four convolutions to isolate the desired convolution and derive the minimum number of zero padding before and after the sequences. The computational complexity of the proposed method is less than that of linear convolution using DFT. In addition, the proposed method gives linear correlation between two sequences when one of the sequences is reversed.

2. PRELIMINARY

Circular convolution as linear convolution and the relation between DCT-2 coefficients and DFT coefficients are described.

Let $x(n)$ be a finite-length sequence of length N such that $x(n) = 0$ outside the range $0 \le n \le N-1$, and let $h(n)$ be a finite-length sequence of length L such that $h(n) = 0$ outside the range $0 \le n \le L-1$.

2.1 Linear convolution

Linear convolution, $y(n)$, of $h(n)$ and $x(n)$ is defined as

$$y(n) = h(n) * x(n) = \sum_{m=-\infty}^{\infty} h(m)x(n-m) \tag{1}$$

where the symbol '*' denotes linear convolution operator. Linear convolution of $h(n)$ and $x(n)$ has the maximum length M, i.e.,

$$M = L + N - 1. \tag{2}$$

2.2 Circular convolution as linear convolution

Let $\tilde{x}(n)$ be a periodic sequence of $x(n)$ with period P as

$$\tilde{x}(n) = \sum_{r=-\infty}^{\infty} x(n-rP) = x(((n))_P) \tag{3}$$

where $((n))_P$ denotes n modulo P, and let $\tilde{h}(n)$ be a periodic sequence of $h(n)$ with period P according to (3).

Circular convolution with period P, $y_P(n)$, of $\tilde{h}(n)$ and $\tilde{x}(n)$ is defined as

$$y_P(n) = \tilde{h}(n) \circledP \tilde{x}(n) = \sum_{m=0}^{P-1} \tilde{h}(m)\tilde{x}(n-m) \tag{4}$$

$$= \sum_{m=0}^{P-1} h(m)x(((n-m))_P), \quad 0 \le n \le P-1. \tag{5}$$

where the symbol 'Ⓟ' denotes circular convolution operator with period P.

The circular convolution $y_P(n)$ of $\tilde{h}(n)$ and $\tilde{x}(n)$ with P-point DFT coefficients $Y(k)$, $H(k)$, and $X(k)$, respectively, is calculated as

$$Y(k) = H(k)X(k), \quad 0 \le k \le P-1 \tag{6}$$

where P-point DFT coefficients, $X(k)$, of $\tilde{x}(n)$ is defined as

$$X(k) = \sum_{n=0}^{P-1} \tilde{x}(n) W_P^{nk}, \quad 0 \le k \le P-1 \tag{7}$$

and W_P denotes $\exp(-j2\pi/P)$.

When the circular convolution has a period of at least M in (2), circular convolution and linear convolution are identical.

2.3 Relation between DCT-2 and DFT

Let $\tilde{x}_s(n)$ be a periodic sequence extended from finite-length sequence $x(n)$ of length N to have period $2N$ as

$$\tilde{x}_s(n) = x(((n))_{2N}) + x(((-n-1))_{2N}). \tag{8}$$

$2N$-point DFT coefficients, $\tilde{X}_s(k)$, of $\tilde{x}_s(n)$ and N-point DCT-2 coefficients of $X_C(k)$ of $\tilde{x}_s(n)$, $0 \le n \le N-1$, i.e., $x(n)$ are related as

$$\tilde{X}_s(k) = \sqrt{2N}/k_k X_C(k) W_{2N}^{-k/2}, \quad 0 \le k \le N-1. \tag{9}$$

where P-point DCT-2 coefficients, $X_C(k)$, of $x(n)$ are defined as

$$X_C(k) = \sqrt{\frac{2}{P}} \sum_{n=0}^{P-1} k_k x(n) \cos\left(\frac{\pi(n+1/2)k}{P}\right), 0 \le k \le P-1 \tag{10}$$

and k_k is the weighting function:

$$k_k = \begin{cases} 1/\sqrt{2}, & k=0 \\ 1, & \text{otherwise} \end{cases}. \tag{11}$$

Circular convolution, $\tilde{y}_{2N}(n)$, of $\tilde{x}_s(n)$ and $\tilde{h}_s(n)$ extended from $x(n)$ and $h(n)$ according to (8) is calculated with their N-point DCT coefficients of $Y_C(k)$, $H_C(k)$, and $X_C(k)$, respectively, as

$$Y_C(k) = H_C(k) X_C(k), \tag{12}$$

and

$$y_C(n) = \sum_{k=0}^{N-1} k_k^2 Y_C(k) \cos\left(\frac{\pi nk}{N}\right). \tag{13}$$

Note that (13) corresponds DCT-1.

Since, from (9), $2N$-point DFT coefficients, $Y(k) = H_s(k)X_s(k)$ is related with $Y_C(k)$ as

$$Y(k) = 2N/(k_k)^2 Y_C(k) W_{2N}^{-k}, \tag{14}$$

The relation between $y_C(n)$ and $\tilde{y}_{2N}(n)$ is expressed by

$$y_C(n) = y_{2N}(n-1), \quad 0 \le n \le N-1. \tag{15}$$

Thus, we can handle DCT-2 coefficients of a sequence as DFT coefficients of the symmetrically extended sequence.

3. PROPOSED COMPUTING METHOD

3.1 Consideration

Circular convolution with period at least the maximum length of a linear convolution is identical to the linear convolution, which can be calculated using DFT. We can also handle DCT-2 coefficients of a sequence as DFT coefficients of the symmetrically extended sequence. Therefore, we analyze circular convolution between symmetrically extended sequences as convolution using DCT.

Let $\tilde{x}_s(n)$ and $\tilde{h}_s(n)$ be periodic sequences with period $2N$ of $x(n)$ and $h(n)$, respectively, according to (8). Circular convolution, $\tilde{y}_{2N}(n)$, with period $2N$ of $\tilde{x}_s(n)$ and $\tilde{h}_s(n)$ is developed from (8) as

$$\begin{aligned}\tilde{y}_{2N}(n) &= \tilde{x}_s(n) \circledast_{2N} \tilde{h}_s(n) \\ &= [x(((n))_{2N}) + x(((-n-1))_{2N})] \\ &\quad \circledast_{2N} [h(((n)_{2N})) + h(((-n-1))_{2N})] \\ &= x(((n))_{2N}) \circledast_{2N} h(((n))_{2N}) \\ &\quad + x(((n))_{2N}) \circledast_{2N} h(((-n-1))_{2N}) \\ &\quad + x(((-n-1))_{2N}) \circledast_{2N} h(((n))_{2N}) \\ &\quad + x(((-n-1))_{2N} \circledast_{2N} h(((-n-1))_{2N}).\end{aligned} \tag{16}$$

The resulting $\tilde{y}_{2N}(n)$ has four circular convolutions with period $2N$. Each circular convolution with period $2N$ is identical to corresponding linear convolution of length M with respect to $x(n)$ and $h(n)$ because $2N > M$.

Our goal is to isolate one linear convolution with respect to $x(n)$ and $h(n)$ from other three in $\tilde{y}_{2N}(n)$.

3.2 Analysis of convolution using $2P$-point DFT

We consider proper zero-padding in sequences as well as linear convolution using DFT.

Without loss of generality, we assume finite-length sequences $x'(n)$ and $h'(n)$ both of length P including $x(n)$ and $h(n)$, respectively, such that $x'(n) = 0$ and $h'(n) = 0$ outside the range $0 \le n \le P-1$, i.e.,

$$x'(n) = \begin{cases} 0, & 0 \le n \le P_1 - 1 \\ x(n-P_1), & P_1 \le n \le P_4 - 1 \\ 0, & P_4 \le n \le P-1 \end{cases} \tag{17}$$

$$h'(n) = \begin{cases} 0, & 0 \le n \le P_2 - 1 \\ h(n-P_2), & P_2 \le n \le P_3 - 1 \\ 0, & P_3 \le n \le P-1 \end{cases} \tag{18}$$

where

$$0 \le P_1 \le P_2 \le P_3 \le P_4 \le P, \tag{19}$$

$$P_3 = P_2 + L, \tag{20}$$

$$P_4 = P_1 + N. \tag{21}$$

Figure 1 illustrates an example of $x(n)$, $h(n)$, $x'(n)$, and $h'(n)$.

From (16), circular convolution, $\tilde{y}'_{2P}(n)$, of periodic sequences $\tilde{x}'_s(n)$ and $\tilde{h}'_s(n)$ extended from $x'(n)$ and $h'(n)$, respectively, to have period $2P$ according to (8) consists of four circular convolutions with period $2P$, i.e.,

$$\begin{aligned}\tilde{y}'_{2P}(n) &= \tilde{x}'_s(n) \circledast_{2P} \tilde{h}'_s(n) \\ &= y_{2P}^{(1)}(n) + y_{2P}^{(2)}(n) + y_{2P}^{(3)}(n) + y_{2P}^{(4)}(n)\end{aligned} \tag{22}$$

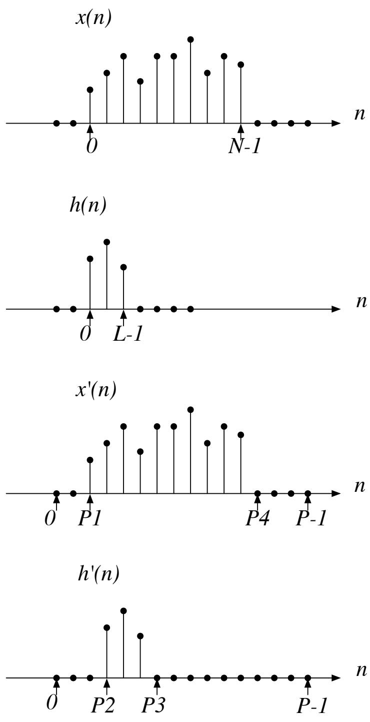


Figure 1: An example of $x(n)$, $h(n)$, $x'(n)$, and $h'(n)$.

where

$$y^{(1)}_{2P}(n) = x'(((n))_{2P}) \textcircled{2P} h'(((n))_{2P}), \quad (23)$$
$$y^{(2)}_{2P}(n) = x'(((n))_{2P}) \textcircled{2P} h'(((-n-1))_{2P}), \quad (24)$$
$$y^{(3)}_{2P}(n) = x'(((-n-1))_{2P}) \textcircled{2P} h'(((n))_{2P}), \quad (25)$$
$$y^{(4)}_{2P}(n) = x'(((-n-1))_{2P}) \textcircled{2P} h'(((-n-1))_{2P}). \quad (26)$$

From (23), (24), (25), and (26), they have symmetry as

$$y^{(1)}_{2P}(n) = y^{(4)}_{2P}(-n-1), \quad (27)$$
$$y^{(2)}_{2P}(n) = y^{(3)}_{2P}(-n-1). \quad (28)$$

The four circular convolutions are identical to corresponding linear convolution with respect to $x(n)$ and $h(n)$ shown in (29), (30), (31), and (32). Figure 2 illustrates the output range of each linear convolution in $\tilde{y}_{2P}(n)$ when $x'(n)$ and $h'(n)$ have enough zero-padding before and after $x(n)$ and $h(n)$, respectively, i.e.,

$$0 \ll P_1 < P_2 < P_3 < P_4 \ll P. \quad (33)$$

Note that $y^{(2)}_{2P}(n)$ and $y^{(3)}_{2P}(n)$ are overlapped. We cannot separate $y^{(2)}_{2P}(n)$ from $y^{(3)}_{2P}(n)$ (See Appendix).

Therefore, we isolate $y^{(1)}_{2P}(n)$ from the other linear convolutions in $\tilde{y}_{2P}(n)$.

3.3 Derivation of proper zero-padding

We derive the locations P_1, P_2, and P for the minimum number of zero-padding to isolate $y^{(1)}_{2P}(n)$.

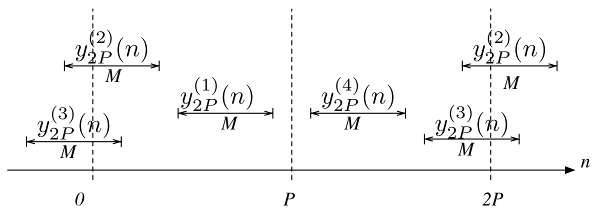


Figure 2: Output range of each linear convolution with respect to $x(n)$ and $h(n)$ in $\tilde{y}'_{2P}(n) = \tilde{h}'_s(n) \textcircled{2P} \tilde{x}'_s(n)$ where $0 \ll P_1 < P_2 < P_3 < P_4 \ll P$.

Firstly, we align the output range of $y^{(2)}_{2P}(n)$ on that of $y^{(3)}_{2P}(n)$ to minimize the period $(2P)$ of $y_{2P}(n)$. Next, we isolate $y^{(1)}_{2P}(n)$ from the others. Since $y^{(1)}_{2P}(n)$ and $y^{(4)}_{2P}(n)$ are symmetry as shown in (27), to isolate $y^{(1)}_{2P}(n)$ from the others is to isolate $y^{(4)}_{2P}(n)$ at the same time. Therefore, the minimum period $(2P)$ of $y_{2P}(n)$ is expressed by

$$2P = 3M. \quad (34)$$

The condition for which the output ranges of $y^{(2)}_{2P}(n)$ and $y^{(3)}_{2P}(n)$ are the same is, from (30) and (31),

$$2P - 1 - P_3 + P_1 = 2P - 1 - P_4 + P_2. \quad (35)$$

It follows that, from (20) and (21),

$$P_2 = P_1 + (N - L)/2. \quad (36)$$

The condition for which $y^{(1)}_{2P}(n)$ is not overlapped on $y^{(3)}_{2P}(n)$ is

$$-2 - P_4 + P_2 + M - 1 < P_1 + P_2. \quad (37)$$

It follows that, from (21) and (36),

$$(L - 3)/2 < P_1. \quad (38)$$

Therefore, the locations P_1, P_2, and P for the minimum number of zero-padding are, from (34), (36), and (38),

$$P_1 > (L - 3)/2, \quad (39)$$
$$P_2 = (N - 3)/2, \quad (40)$$
$$P > (3M)/2. \quad (41)$$

Thus, proper zero-padded sequences $x'(n)$ and $h'(n)$ according to (39), (40), and (41) are generated to isolate $y^{(1)}_{2P}(n)$. The linear convolutions can be obtained from $x'(n)$ and $h'(n)$ with P-point DCT-2 in the proposed method. Note that the number of zero-padding can be more reduced if the transient output in the resulting is not required. Details are omitted due to space limitation.

3.4 Steps of the proposed method

Steps of the proposed method is summarized as follows:

1. Make zero-padded signals $x'(n)$ and $h'(n)$ from $x(n)$ and $h(n)$ according to (39), (40), and (41).
2. Apply DCT-2 to $x'(n)$ and $h'(n)$.
3. Multiply the DCT-2 coefficients element by element.
4. Apply DCT-1 to the product.
5. Extract the output.

$$y_{2P}^{(1)}(n) = \begin{cases} x(n) * h(n), & P_1 + P_2 \leq n \leq P_1 + P_2 + M - 1 \\ 0, & \text{otherwise} \end{cases} \tag{29}$$

$$y_{2P}^{(2)}(n) = \begin{cases} x(n) * h(-n-1), & 2P - 1 - P_3 + P_1 \leq n \leq 2P - 1 - P_3 + P_1 + M - 1 \\ 0, & \text{otherwise} \end{cases} \tag{30}$$

$$y_{2P}^{(3)}(n) = \begin{cases} x(-n-1) * h(n), & 2P - 1 - P_4 + P_2 \leq n \leq 2P - 1 - P_4 + P_2 + M - 1 \\ 0, & \text{otherwise} \end{cases} \tag{31}$$

$$y_{2P}^{(4)}(n) = \begin{cases} x(-n-1) * h(-n-1), & 2P - 1 - P_3 + 2P - 1 - P_4 \leq n \leq 2P - 1 - P_3 + 2P - 1 - P_4 + M - 1 \\ 0, & \text{otherwise} \end{cases} \tag{32}$$

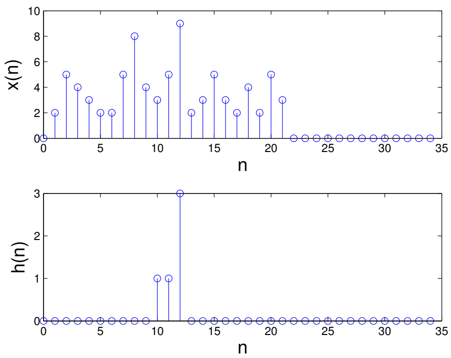


Figure 3: Zero-padded sequences $x'(n)$ and $h'(n)$ with $P_1 = 1$, $P_2 = 10$, $P = 35$ according to (39), (40), and (41).

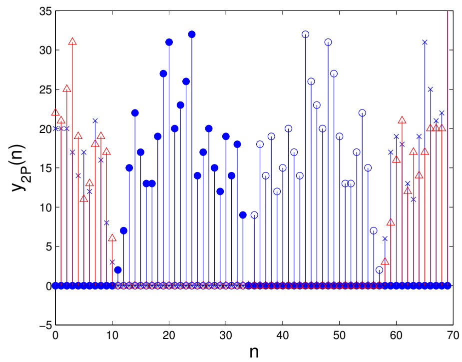


Figure 4: Four linear convolutions with respect to $x(n)$ and $h(n)$ by $2P$-point DFT where $P_1 = 1$, $P_2 = 10$, $P = 35$. Symbols '•', '×', '△', and '○' denote $y_{2P}^{(1)}(n)$, $y_{2P}^{(2)}(n)$, $y_{2P}^{(3)}(n)$, and $y_{2P}^{(4)}(n)$, respectively. $y_{2P}^{(1)}(n)$ can be isolated from the others due to proper zero-padding.

3.5 Linear correlation

We can also consider $y_{2P}^{(2)}(n)$ and $y_{2P}^{(3)}(n)$ as linear correlation with respect to $x(n)$ and $h(n)$ from (24) and (25). When setting $x'(n)$ in (17) to have the reverse of $x(n)$, we can obtain the linear correlation with respect to $x(n)$ and $h(n)$ from $y_{2P}(n)$ of $x'(n)$ with $h'(n)$.

4. SIMULATION

4.1 Experimental result

We performed linear convolution of $x(n)$ and $h(n)$ of length $N = 21$ and $L = 3$, respectively.

Figure 3 shows zero-padded sequences, $x'(n)$ and $y'(n)$ with $P_1 = 1$, $P_2 = 10$, $P = 35$, according to (39), (40), and (41). Figure 4 shows four convolutions using $2P$-point DFT when $2P$-point sequences are generated from $x'(n)$ and $y'(n)$ shown in 3. Symbols '•', '×', '△', and '○' denote $y_{2P}^{(1)}(n)$, $y_{2P}^{(2)}(n)$, $y_{2P}^{(3)}(n)$, and $y_{2P}^{(4)}(n)$, respectively. We can see that $y_{2P}^{(1)}(n)$ is properly isolated from the others. Figure 5 shows the resulting convolution using the proposed method. Linear convolution of $x(n)$ and $h(n)$ can be obtained from P-point resulting.

4.2 Computational complexity

Computational complexity of the proposed method is compared to that of linear convolution using DFT.

One of the advantages of DCT over DFT is calculation with real numbers. A fast algorithm of DCT for a sequence of length N runs with

$$M_{real} = (N/2)\log_2 N + 1, \tag{42}$$
$$A_{real} = (3N/2)\log_2 N - N + 1 \tag{43}$$

where M_{real} and A_{real} denote the number of multiplications and additions with real numbers, respectively [6]. FFT, the fast algorithm of DFT, on the other hand, for a sequence of length N runs with

$$M_{complex} = (N/2)\log_2 N, \tag{44}$$
$$A_{complex} = N\log_2 N \tag{45}$$

where $M_{complex}$ and $A_{complex}$ are the number of multiplications and additions with complex numbers, respectively [7].

Figures 6(a) and 6(b) show the number of multiplications and additions in real numbers, respectively, for linear convolution of length M, in which one multiplication in complex numbers can be realized by three multiplications and three additions in real numbers by Nakayama's method [8]. The number of operations in the proposed method is less than that in linear convolution using FFT.

5. CONCLUSION

We have proposed a computing method for linear convolution using DCT. From the relation between DCT-2 and DFT

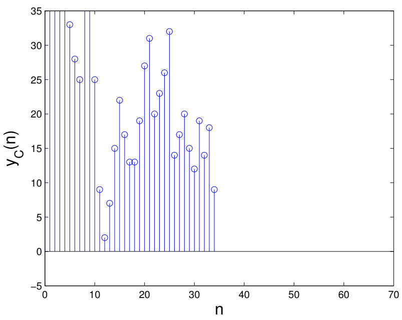


Figure 5: Linear convolution using the proposed method. Linear convolution, $y_{2P}^{(1)}(n)$, can be isolated from the others due to proper zero-padding using P-point DCT coefficients.

coefficients, we have analyzed circular convolution between sequences extended symmetrically for DCT-2. The output range of linear convolution has been analyzed and the minimum number of zero-padding before and after sequences has been derived. We have shown the correctness and effectiveness of the proposed method in some simulations comparing to convolution using DFT.

REFERENCES

[1] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-time Signal Processing,* Prentice Hall, 1999

[2] K.R.Rao and P.Yip, *Discrete Cosine Transform*, Academic Press. Inc., 1990

[3] S. A. Martucci, "Symmetric convolution and the discrete sine and cosine transforms," *IEEE Trans. Signal Process.*, vol.42, no.5, May 1994

[4] V. G. Reju, S. N. Koh, and I. Y. Soon, "Convolution using discrete sine and cosine transforms," *IEEE Signal Processing Letters*, vol.14, no.7, July 2007

[5] K.Suresh, T. V. Sreenivas, "Linear filtering in DCT IV/DST IV and MDCT/MDST domain," *EURASIP, Signal Processing*, vol.89 pp.1081–1089, 2009

[6] Z. Wang, "On computing the discrete Fourier and cosine transforms," *IEEE Trans. Accoust., Speech, Signal Process.*, vol.33, no.4, pp.1341-1344, 1985

[7] H. J. Nussbaumer, *Fast Fourier transform and convolution algorithms,* Springer-Verlog, 1982

[8] K. Nakayama, "A new discrete Fourier transform algorithm using butterfly structure fast convolution," *IEEE Trans. Acoust. Speech Signal Process.*, vol.ASSP-33, no.4, pp.1197–1208, Oct. 1985

Appendix

We show that we cannot separate $y_{2P}^{(2)}(n)$ from $y_{2P}^{(3)}(n)$.

Firstly, we assume

$$y_{2P}^{(2)}(n) < y_{2P}^{(3)}(n). \tag{46}$$

That is, the right side of output range of $y_{2P}^{(2)}(n)$ is less than the left side of the output range of $y_{2P}^{(3)}(n)$:

$$2P-1-P_3+P_1+M-1 < 2P-1-P_4+P_2. \tag{47}$$

It follows that

$$P_1-P_2 < -(N-1). \tag{48}$$

If $P_1 = P_2$, then (48) cannot be satisfied because the right hand is zero and the left hand is negative. If $P_1 < P_2$, then (48) cannot be satisfied because $P_2 \leq N-1$ under $0 \leq P_1 < P_2$. Hence, (46) is invalid.

Next, we assume

$$y_{2P}^{(3)}(n) < y_{2P}^{(2)}(n). \tag{49}$$

That is, the right side of output range of $y_{2P}^{(3)}(n)$ is less than the left side of the output range of $y_{2P}^{(2)}(n)$:

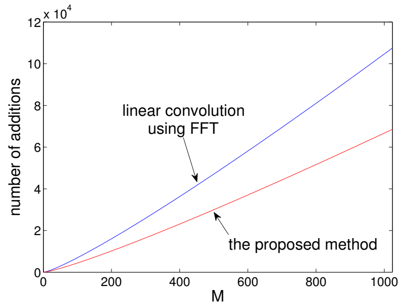
$$2P-1-P_4+P_2+M-1 < 2P-1-P_3+P_1. \tag{50}$$

It follows that

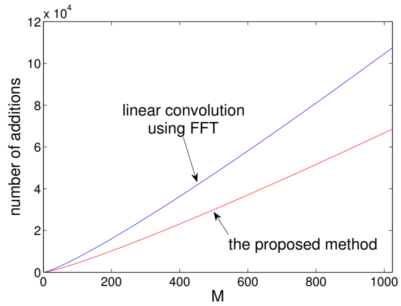
$$P_2-P_1 < -L+1. \tag{51}$$

If $P_1 = P_2$, then (51) cannot be satisfied because $-L+1 \leq 0$. If $P_1 < P_2$, then (51) cannot be satisfied because the right hand is positive and the left hand is $-L+1 \leq 0$. Hence, (49) is invalid.

Therefore, $y_{2P}^{(2)}(n)$ cannot be separated from $y_{2P}^{(3)}(n)$.



(a) The number of multiplications in real numbers.



(b) The number of additions in real numbers.

Figure 6: The number of multiplications in real numbers for linear convolution of length M. One multiplication in complex numbers is three multiplications and three additions in real numbers by Nakayama's method.