Optimization of two-dimensional real convolutions

Richard E. Crandall and James Farrell
Center for Advanced Computation, Reed College, Portland OR

Abstract. Cyclic convolution of two two-dimensional signals, each of dimension M-by-N, can be performed by traditional Fourier methods in $O(MN \log MN)$ complex operations. When the signals are real, however, so that the final convolution is likewise real, there is no need to involve imaginary numbers at any stage of the calculation. Instead, one may use either the real-valued fast Fourier transform (RVFFT, also known as the Hermitian FFT), or the fast Hartley transform (FHT). Then the necessary spectral (dyadic) multiply of the classical convolution theorem becomes a different but related operation amongst RVFFT or FHT coefficients. By establishing the precise formulæ for the new spectral multiplies, we establish herein a two-dimensional real convolution scheme that is fast, and also in-place: no memory substantially beyond that of the input signals is required.

Dec 1997

1. Definitions and terminology

This paper concerns cyclic convolutions of two-dimensional real signals. We presume that a signal x comprises elements x_{jk}, where coordinate j can range from 0 to $M-1$, k from 0 to $N-1$. Given such an array, we define the following discrete transformations, where a and b range over the same respective intervals:

1. The Fourier transform (FFT) $\hat{x}$:

$$\hat{x}_{ab} = \sum_{k=0}^{N-1} \sum_{j=0}^{M-1} x_{jk} \ \exp\left(-2\pi i j a/M\right) \exp\left(-2\pi i k b/N\right). \tag{1.1}$$

2. The Hermitian, or real-valued Fourier transform (RVFFT) $\hat{X}$:

$$\hat{X}_{ab} = \sum_{k=0}^{N-1} \sum_{j=0}^{M-1} x_{jk} \ \mathrm{CS}_{a,M}\left(\frac{2\pi j a}{M}\right) \mathrm{CS}_{b,N}\left(\frac{2\pi k b}{N}\right), \tag{1.2}$$

where

$$\mathrm{CS}_{c,P}(x) = \begin{cases} \cos(x) & 0 \le c \le P/2 \\ \sin(x) & P/2 < c < P \end{cases}. \tag{1.3}$$

3. The Hartley transform (FHT) $\hat{H}$:

$$\hat{H}_{ab} = \sum_{k=0}^{N-1} \sum_{j=0}^{M-1} x_{jk} \ \mathrm{cas}\left(\frac{2\pi j a}{M}\right) \mathrm{cas}\left(\frac{2\pi k b}{N}\right), \tag{1.4}$$

where $\mathrm{cas}(x) = \cos(x) + \sin(x)$.

Fast algorithms ($O(N \log N)$) to compute each of these transforms are known.

Provided the original 2-dimensional signal x is real, each of the three transforms contains equivalent information: given any one of the three transforms $\hat{x}$, $\hat{X}$, $\hat{H}$, one may determine unambiguously the other two.

2. General strategy

Under the FFT, the cyclic convolution $x \otimes y$ assumes the form of simple dyadic multiplication. That is, the classical convolution theorem can be expressed thus:

$$(\hat{x} \otimes \hat{y})_{ab} = \hat{x}_{ab} \ \hat{y}_{ab}. \tag{2.1}$$

Convolution under the RVFFT and FHT is more complicated, but can be derived by exploiting their relation to the DFT.

For notational simplicity, we shall abbreviate $\cos(2\pi j a/M)$ as C_a and define C_b, S_a, and S_b similarly. Furthermore, we define

$$\delta_a^+ = \begin{cases} 1 & 0 \le a \le M/2 \\ 0 & M/2 < a < M \end{cases}, \tag{2.2}$$

with δ_b definited similarly and $\delta^- = 1 - \delta^+$. We can then write our three transforms as

$$\begin{array}{lccccccccc}
\hat{x}_{ab} = \sum_{j,k} & & \mathrm{C}_a\mathrm{C}_b- & i & \mathrm{C}_a\mathrm{S}_b- & i & \mathrm{S}_a\mathrm{C}_b+ & & \mathrm{S}_a\mathrm{S}_b \\
\hat{X}_{ab} = \sum_{j,k} & (\delta_a^+\delta_b^+) & \mathrm{C}_a C_b+ & (\delta_a^+\delta_b^-) & \mathrm{C}_a\mathrm{S}_b+ & (\delta_a^-\delta_b^+) & \mathrm{S}_a\mathrm{C}_b+ & (\delta_a^-\delta_b^-) & \mathrm{S}_a\mathrm{S}_b \\
\hat{H}_{ab} = \sum_{j,k} & & \mathrm{C}_a\mathrm{C}_b+ & & \mathrm{C}_a\mathrm{S}_b+ & & \mathrm{S}_a\mathrm{C}_b+ & & \mathrm{S}_a\mathrm{S}_b
\end{array} \tag{2.3}$$

Now, we pick two relations (the Fourier and, for purposes of argument, the Hermitian Fourier) and treat them as a system of two equations in four variables to find $\hat{x}$ in terms of $\hat{X}$ or *vice versa.* If

$$\hat{x} = F(\hat{X}) \text{ and } \hat{X} = F^{-1}(\hat{x}), \tag{2.4}$$

then

$$\begin{aligned} \hat{X} \otimes \hat{Y} &= F^{-1}(\hat{x} \otimes \hat{y}) \\ &= F^{-1}[F(\hat{X})F(\hat{Y})]. \end{aligned} \tag{2.5}$$

To determine the convolution relation, then, we need only demonstrate the relation between $\hat{x}_{ab}$ and $\hat{X}_{ab}$ for a *single* array x_{jk}.

3. Convolution under the Hermitian Fourier transform

Equation (1) cannot be solved in the general case for the RVFFT. It is therefore necessary to divide the array x_{jk} into several regions:

Region 1: The four points $(0,0)$, $(M/2,0)$, $(0,N/2)$, $(M/2,N/2)$.

In this region, the two transforms assume identical values and the convolution is straightforward:

$$\begin{array}{ll} \hat{x}_{0,0} = \hat{X}_{0,0} & \hat{x}_{M/2,0} = \hat{X}_{M/2,0} \\ \hat{x}_{0,N/2} = \hat{X}_{0,N/2} & \hat{x}_{M/2,N/2} = \hat{X}_{M/2,N/2}, \end{array} \tag{3.1}$$

resulting, for the convolution $\hat{Z} = \hat{X} \otimes \hat{Y}$ in simple dyadic multiplication:

$$\hat{Z}_{ab} = \hat{X}_{ab}\hat{Y}_{ab}, \qquad a = 0, M/2, \quad b = 0, N/2. \tag{3.2}$$

Region 2: The four lines where $a = 0, M/2$ or $b = 0, N/2$, excluding the Region 1.

In this region, the relations are defined as follows:

$$\begin{aligned} \hat{x}_{ab} &= \hat{x}^*_{-a,-b} = \hat{X}_{ab} + i\hat{X}_{-a,-b} \\ \hat{X}_{ab} &= \begin{cases} \delta_b^+\mathrm{Re}(\hat{x}_{ab}) - \delta_b^-\mathrm{Im}(\hat{x}_{ab}) & a = 0, M/2 \\ \delta_a^+\mathrm{Re}(\hat{x}_{ab}) - \delta_a^-\mathrm{Im}(\hat{x}_{ab}) & b = 0, N/2, \end{cases} \end{aligned} \tag{3.3}$$

where $-a = M - a$ and $-b = N - b$ except for $a, b = 0$ to bring the index within the required range of values. This results in the convolution relation

$$\begin{array}{ll} \left.\begin{aligned} \hat{Z}_{ab} &= \hat{X}_{ab}\hat{Y}_{ab} - \hat{X}_{a,-b}\hat{Y}_{a,-b} \\ \hat{Z}_{a,-b} &= \hat{X}_{a,-b}\hat{Y}_{ab} + \hat{X}_{ab}\hat{Y}_{a,-b} \end{aligned}\right\} & a = 0, M/2, \quad 0 < b < N/2 \\ \left.\begin{aligned} \hat{Z}_{ab} &= \hat{X}_{ab}\hat{Y}_{ab} - \hat{X}_{-a,b}\hat{Y}_{-a,b} \\ \hat{Z}_{-a,b} &= \hat{X}_{-a,b}\hat{Y}_{ab} + \hat{X}_{ab}\hat{Y}_{-a,b} \end{aligned}\right\} & a = 0 < a < M/2, \quad b = 0, N/2. \end{array} \tag{3.4}$$

Region 3: The remaining array elements, consisting of four quadrants: $[1, M/2-1] \times [1, N/2-1]$ and its reflection about $a = M/2$, $b = N/2$, and both.

In this case, if we simply specify $a, b \in [1, M/2-1] \times [1, N/2-1]$, the values in all four quadrants will emerge at once:

$$\begin{aligned}
\hat{x}_{ab} = \hat{x}^*_{-a,-b} &= \hat{X}_{ab} - i\hat{X}_{a,-b} - i\hat{X}_{-a,b} - \hat{X}_{-a,-b} \\
\hat{x}_{a,-b} = \hat{x}^*_{-a,b} &= \hat{X}_{ab} + i\hat{X}_{a,-b} - i\hat{X}_{-a,b} + \hat{X}_{-a,-b} \\
2\hat{X}_{ab} &= \mathrm{Re}(\hat{x}_{ab} + \hat{x}_{-a,b}) \\
2\hat{X}_{-a,b} &= \mathrm{Im}(\hat{x}_{-a,b} - \hat{x}_{ab}) \\
2\hat{X}_{a,-b} &= -\mathrm{Im}(\hat{x}_{ab} + \hat{x}_{-a,b}) \\
2\hat{X}_{-a,-b} &= \mathrm{Re}(\hat{x}_{-a,b} - \hat{x}_{ab}).
\end{aligned} \tag{3.5}$$

In order to obtain the convolution relation for the four quadrants, define the quantities A, B, C, D as follows:

$$\begin{aligned}
A &= (\hat{X}_{ab} - \hat{X}_{-a,-b})(\hat{Y}_{ab} - \hat{Y}_{-a,-b}) - (\hat{X}_{a,-b} + \hat{X}_{-a,b})(\hat{Y}_{a,-b} + \hat{Y}_{-a,b}) \\
B &= (\hat{X}_{ab} + \hat{X}_{-a,-b})(\hat{Y}_{ab} + \hat{Y}_{-a,-b}) - (\hat{X}_{a,-b} - \hat{X}_{-a,b})(\hat{Y}_{a,-b} - \hat{Y}_{-a,b}) \\
C &= (\hat{X}_{a,-b} + \hat{X}_{-a,b})(\hat{Y}_{ab} - \hat{Y}_{-a,-b}) + (\hat{X}_{ab} - \hat{X}_{-a,-b})(\hat{Y}_{a,-b} + \hat{Y}_{-a,b}) \\
D &= (\hat{X}_{a,-b} - \hat{X}_{-a,b})(\hat{Y}_{ab} + \hat{Y}_{-a,-b}) + (\hat{X}_{ab} + \hat{X}_{-a,-b})(\hat{Y}_{a,-b} - \hat{Y}_{-a,b}).
\end{aligned} \tag{3.6}$$

Then

$$\begin{aligned}
2\hat{Z}_{ab} &= A + B \\
2\hat{Z}_{-a,-b} &= B - A \\
2\hat{Z}_{a,-b} &= C + D \\
2\hat{Z}_{-a,b} &= C - D.
\end{aligned} \tag{3.7}$$

4. Convolution under the Hartley transform

In the case of Hartley convolution, the solution to (1) is much easier due to the relationship between $\exp(ix)$ and $\mathrm{cas}(x)$ which does not depend on the conditionally-valued δ's. Upon close examination the Hartley and Fourier transforms may be written in terms of each other as follows:

$$\begin{aligned}
2\hat{H}_{ab} &= (i\hat{x}_{ab} + \hat{x}_{-a,b} + \hat{x}_{a,-b} - i\hat{x}_{-a,-b}) \\
2\hat{x}_{ab} &= (-i\hat{H}_{ab} + \hat{H}_{-a,b} + \hat{H}_{a,-b} + i\hat{H}_{-a,-b})
\end{aligned} \tag{4.1}$$

Applying the method above for $\hat{L} = \hat{H} \otimes \hat{K}$, the Hartley analog of $\hat{z} = \hat{x} \otimes \hat{y} = \hat{x}\hat{y}$, we write

$$2\hat{L}_{ab} = (i\hat{z}_{ab} + \hat{z}_{-a,b} + \hat{z}_{a,-b} - i\hat{z}_{-a,-b}) \tag{4.2}$$

After substitution for $\hat{z}$ in this expression and algebraic manipulation that is lengthy but straightforward, we obtain the final relation

$$\begin{aligned} 4\hat{L}_{ab} &= (\hat{H}_{ab} + \hat{H}_{-a,-b})(\hat{K}_{ab} + \hat{K}_{-a,-b}) \\ &- (\hat{H}_{a,-b} - \hat{H}_{-a,b})(\hat{K}_{a,-b} - \hat{K}_{-a,b}) \\ &- (\hat{H}_{-a,-b} - \hat{H}_{ab})(\hat{K}_{-a,b} + \hat{K}_{a,-b}) \\ &- (\hat{H}_{-a,b} + \hat{H}_{a,-b})(\hat{K}_{-a,-b} - \hat{K}_{ab}) \end{aligned} \tag{4.3}$$

Unlike the Hermitian relation, this expression does not require division of x_{jk} into regions.

6. Algorithm

Input: Two real M-by-N arrays x, y.
Ouput: The cyclic convolution $z = x \otimes y$
Complexity: $O(MN \log MN)$ real multiplications.

1) Compute the in-place 2D RVFFTs or FHTs of x, y to give $\hat{X}, \hat{Y}$ or $\hat{H}, \hat{K}$. That is, compute the 1D transforms of all the rows followed by the 1D transforms of all the columns.

2) Perform the appropriate dyadic multiplications. For the RVFFT, use (3.2), (3.4), and (3.7) on the appropriate regions of the transformed arrays. For the FHT, use (4.3) for the entire arrays. That is, compute

$$\hat{Z} = \hat{X} \otimes \hat{Y} \tag{6.1}$$

or

$$\hat{L} = \hat{H} \otimes \hat{K} \tag{6.2}$$

3) Compute the inverse RVFFT or FHT to give z:

$$z = \mathrm{RVFFT}^{-1}(\hat{Z}) = \mathrm{FHT}^{-1}(\hat{L}). \tag{6.3}$$

For optimal memory efficiency, the values for $\hat{Z}$ or $\hat{L}$ can be computed in place by overwriting the values for $\hat{Y}$ or $\hat{K}$.

6. Speed and memory

As with standard 2-dimensional Fourier transforms, one can calculate either the RVFFT or the FHT by doing all rows in place, then all columns in place. After doing thus for each of two signals x, y to be convolved, one invokes the dyadic multiplication of the appropriate form. The key to memory efficiency is: this dyadic multiply can be done "in-place." That is, the 2-dimensional transform of y can be used to store the dyadic product, term by term. (One may access four elements of y's transform at a time, and the corresponding four of x's transform, then lay these back down after the multiplication algebra, onto the y array.) Finally, one does all columns of y in-place, then the rows; whence y has become the required cyclic convolution.

We tested implementations of the RVFFT and Hartley schemes on a DEC Alpha CPU. The RVFFT convolution was always slightly faster, though the Hartley, in being somewhat more memory-cache friendly, might perform better on other machinery. The standard DFT is between two and three times slower than either RVFFT or FHT and uses twice the memory to store the redundant imaginary parts of the arrays.

A commercial implementation of the schemes described herein is available as the software release

`science2/transfor/fft`

from Perfectly Scientific, Inc., Portland OR, located at

`http://www.perfsci.com/`.