

Detection of Linear and Cubic Interpolation in JPEG Compressed Images

Andrew C. Gallagher

Eastman Kodak Company

andrew.gallagher@kodak.com

Abstract

A novel algorithm is introduced that can detect the presence of interpolation in images prior to compression as well as estimate the interpolation factor. The interpolation detection algorithm exploits a periodicity in the second derivative signal of interpolated images. The algorithm performs well for a wide variety of interpolation factors, both integer factors and non-integer factors. The algorithm performance is noted with respect to a digital camera's "digital zoom" feature. Overall, the algorithm has demonstrated robust results and might prove to be useful for situations where an original resolution of the image determines the action of an image processing chain.

1. Introduction

Interpolation is the process of estimating the value of a signal at positions intermediate to the original samples. Generally, this is accomplished by fitting a continuous function to the known samples and evaluating the function at the desired locations. In order to avoid aliasing with no degradation of valid signal, an ideal low-pass filter must be used. In the spatial domain, this filter is represented as a sinc function [1], which is infinite in spatial extent. In practice, the windowed sinc function is large and often produces ringing artifacts because of Gibbs phenomenon. A linear low-pass filter or a cubic low-pass filter [2], shown in Figure 1 is commonly used in image processing as a useful alternative to the sinc function. The equations for the linear and cubic interpolation filters $h_l(x)$ and $h_c(x)$ are:

$$h_l(x) = 1 - 2\left|x - \frac{1}{2}\right| \quad \text{for } |x| \leq 1 \quad (1)$$

$$h_c(x) = \frac{3}{2}|x^3| - \frac{5}{2}x^2 + 1 \quad \text{for } |x| \leq 1 \quad (2)$$

$$h_c(x) = -\frac{1}{2}|x^3| + \frac{5}{2}x^2 - 4|x| + 2 \quad \text{for } 1 < |x| \leq 2$$

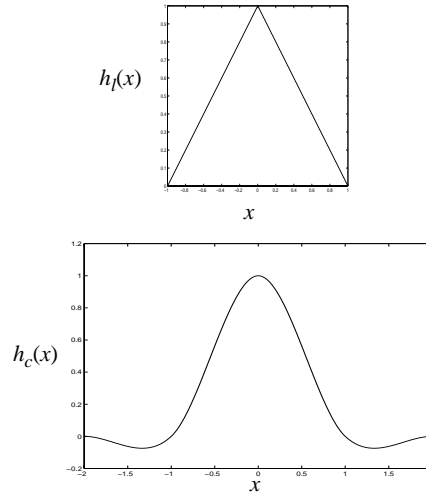


Figure 1: The interpolation function for linear interpolation (top) and cubic interpolation (bottom).

Linear interpolation uses up to two samples, and cubic interpolation uses up to four samples of an original signal to calculate the value of an interpolated sample. In general, it is agreed that the cubic interpolation filter yields superior results to the bilinear interpolation filter.

An interpolated image contains clues that can be extracted to determine the image's history. In fact, the interpolated image can be manipulated to produce a signal that is periodic with the period equal to the resampling rate N . It will be shown that an unknown signal can be analyzed to determine not only if it was interpolated from some original signal, but also the interpolation factor can be determined.

2. Analysis of interpolated image signals

Referring to Figure 2, a one-dimensional digital signal (i.e., an image row or column) is analyzed. Suppose an interpolation function is $h(x)$ and an original signal is $y(n)$, where n is an integer. The original signal is interpo-

lated to produce sample $i(x_0)$, where x_0 is a real number. Therefore $i(x_0)$ can be written as a matrix product:

$$i(x_0) = \mathbf{y}^T \mathbf{h} \quad (3)$$

where matrix \mathbf{y} has entries of $y_k = y(n_0 - k)$, and n_0 is the greatest integer not greater than x_0 . The filter matrix \mathbf{h} has entries $h_k = h(k - \delta)$, where $\delta = x_0 - n_0$.

Equation (3) can be re-written as:

$$i(x_0) = \sum_{k=-\infty}^{\infty} y_k h_k \quad (4)$$

Consider the samples in the interpolated signal that precede and follow $i(x_0)$. These samples are given as $i(x_0 - \Delta)$ and $i(x_0 + \Delta)$, respectively. The interval between samples in the interpolated signal is Δ , where $\Delta = \frac{1}{N}$.

From these samples, a second derivative can be computed:

$$s(x_0) = 2i(x_0) - i(x_0 - \Delta) - i(x_0 + \Delta) \quad (5)$$

This signal can also be written as a matrix product:

$$s(x_0) = \mathbf{y}^T \mathbf{c}(\delta) \quad (6)$$

where the entries of matrix \mathbf{c} are given as:

$$c_k(\delta) = 2h(k - \delta) - h(k - \delta + \Delta) - h(k - \delta - \Delta)$$

The variance $v(x_0)$ of the signal $s(x_0)$ as a function of position x_0 can be computed as:

$$v(x_0) = \mathbf{c}(\delta)^T \mathbf{K} \mathbf{c}(\delta) \quad (7)$$

where \mathbf{K} is the covariance matrix of the samples of the original signal $y(n)$.

If we assume the original pixel values are samples from a Gaussian distribution with variance 1 and mean 0, the

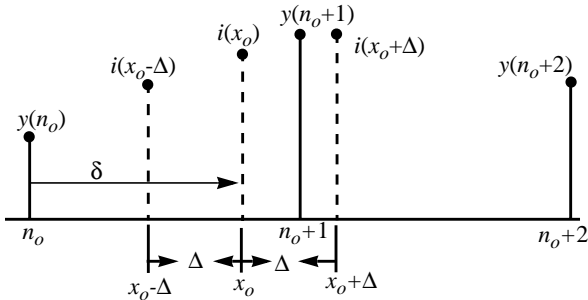


Figure 2: An original signal with samples $y(n_0)$, $y(n_0 + 1)$, and $y(n_0 + 2)$ is interpolated to produce samples $i(x_0 - \Delta)$, $i(x_0)$, and $i(x_0 + \Delta)$.

covariance matrix \mathbf{K} is the identity, and the variance becomes:

$$v(x_0) = \sum_{k=-\infty}^{\infty} c_k(\delta)^2 \quad (8)$$

Notice that the entries of the coefficient matrix \mathbf{c} have the relationship:

$$c_k(\delta + 1) = c_{k-1}(\delta) \quad (9)$$

Therefore,

$$v(x_0) = v(x_0 + 1) \quad (10)$$

so $v(x_0)$ is periodic over x_0 , with period 1. The conclusion is that the variance of the second derivative of an interpolated signal has a periodicity equal to the sampling rate of the original signal. Practically, this characteristic can be exploited to determine whether a given image has been created by a low-order interpolator and the rate of interpolation N .

Suppose samples of an interpolated signal $i(x)$ are used to form a new sequence $i(m)$

$$i(m) = i\left(\frac{x}{\Delta}\right) \quad (11)$$

where m is an integer. We can compute the signal $v(x_0)$ from signal $i(m)$ at integer locations. In an image, this variance signal can be estimated because there are many rows (or columns) over which statistics can be estimated. Based on (9), we expect that the periodicity of the signal $v(x_0)$ will be $\frac{1}{\Delta} = N$, the resampling factor.

To satisfy the Nyquist criteria, the $v(x_0)$ signal must be sampled at least twice per period. This is satisfied only when $\Delta \leq 0.5$ (i.e., the resampling rate N is at least 2.0). Lower resampling rates will result in aliasing.

In the case where the interpolation function is the linear interpolation filter (1), the coefficient matrix \mathbf{c} can be shown as:

$$\begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ when } \delta \geq \Delta \text{ and } \delta + \Delta \leq 1 \quad (12)$$

and

$$\begin{bmatrix} c_{-1} \\ c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} \delta - \Delta \\ -2(\delta - \Delta) \\ \delta - \Delta \end{bmatrix} \text{ when } \delta > \Delta \quad (13)$$

The resulting periodic function $v(x_0)$ is shown in Figure 3.

When the interpolating function is the cubic filter (2), the analysis of the cubic interpolator becomes more involved. When $\delta \geq \Delta$ and $\delta + \Delta \leq 1$:

$$\begin{bmatrix} c_{-1} \\ c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} \Delta^2(2-3\delta) \\ \Delta^2(-5+9\delta) \\ \Delta^2(4-9\delta) \\ \Delta^2(-1+3\delta-1+3) \\ 0 \end{bmatrix}$$

when $\delta + \Delta > 1$:

$$\begin{bmatrix} c_{-1} \\ c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} (\Delta^3 + \Delta + \delta^3 + \delta + 3\delta\Delta(\delta - \Delta))/2 + \Delta^2 - \delta^2 - 2\delta\Delta \\ 1 - 2\Delta^3 - 4\Delta - 2\delta^3 + 5\delta^2 - 4\delta + 10\delta\Delta + 3\delta\Delta^2 - 6\delta^2\Delta \\ 3\Delta^3 - 5\Delta^2 + 9\Delta + 3\delta^3 - 9\delta^2 + 9\delta - 18\delta\Delta + 9\delta^2\Delta - 3 \\ 3 + \Delta(6\Delta - 2\Delta^2 - 8) + \delta(7\delta - 2\delta^2 - 8) + \delta\Delta(14 - 6\delta - 3\Delta) \\ (\Delta + \delta - 1)^2(\Delta + \delta - 2)/2 \end{bmatrix}$$

The resulting periodic function $v(x_0)$ is shown in Figure 4.

It has been proven that the variance of the second derivative signal is periodic with the period equal to the resampling factor N . This observation can be used as the underlying model for an algorithm that detects whether a digital image has been interpolated and determines the interpolation rate N .

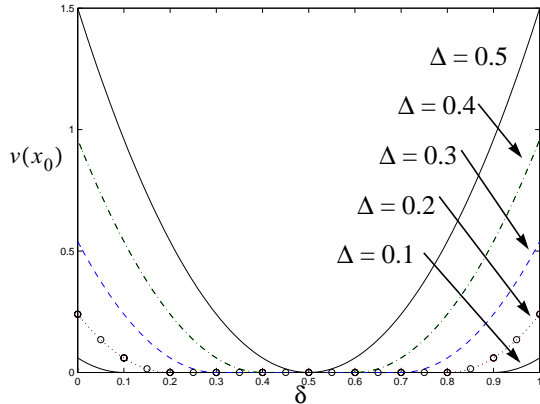


Figure 3: The resulting $v(x_0)$ signal for the linear interpolation filter over the range $0 \leq \delta \leq 1$ (one period) for a variety of sampling intervals Δ . In all cases, the $v(x_0)$ signal is maximized at $\delta = 0$ and minimized for $\delta = \frac{1}{2}$.

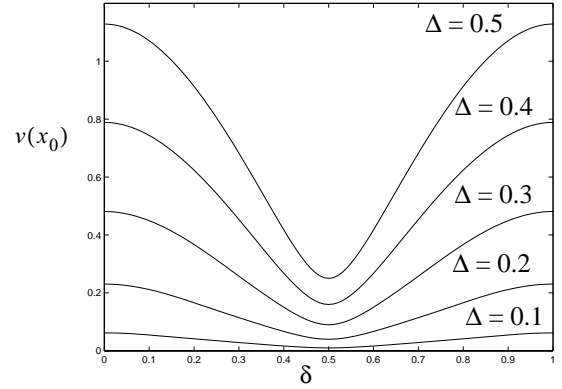


Figure 4: The resulting $v(x_0)$ signal for the cubic interpolation filter over the range $0 \leq \delta \leq 1$ (one period) for a variety of sampling intervals Δ . In all cases, the $v(x_0)$ signal is maximized at $\delta = 0$ and minimized for $\delta = \frac{1}{2}$.

3. Interpolation detection algorithm

The interpolation detection algorithm has the goal of determining if an image of unknown origin has been produced with a low-order interpolation such as bilinear or bicubic interpolation. The interpolation detection algorithm has the additional goal of determining the factor of interpolation when the image has been interpolated.

A block diagram of an algorithm designed to detect interpolation in a digital image is shown in Figure 5. The output of the algorithm is \hat{N} , the estimate of the interpola-

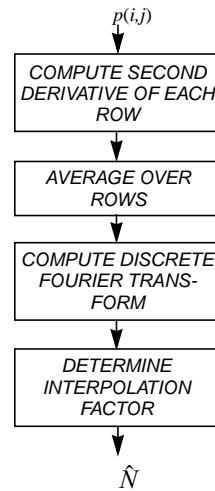


Figure 5: Block diagram of the interpolation detection algorithm.

tion factor N . The algorithm functions by first computing the second derivative of each row of the image. Next, the absolute values of each second derivative row are averaged together in order to obtain a mean second derivative trace. This signal is proportional to the variance signal (8). If an image has been interpolated, this trace will exhibit a periodicity related to the interpolation rate. Finally, the discrete Fourier transform (DFT) of the mean second derivative signal is examined for peaks, and the corresponding frequencies of the peak locations are used to determine \hat{N} , which is related to the period corresponding to the frequency at a peak in the DFT.

3.1. Differentiation

The first step of the interpolation detection algorithm is to compute the second derivative of each row of the input image. Alternatively, whereas the interpolation algorithms we are considering have similar operations with respect to rows and columns, the derivative of each column could be computed. It is assumed that the image input to the algorithm is represented as $p(i,j)$, where $0 \leq i < R$ and $0 \leq j < C$, R is the number of rows in the image, and C is the number of columns in the image.

The second derivative of each row is computed for $1 \leq j < C - 1$ with the following difference equation:

$$s_p(i, j) = 2p(i, j) - p(i, j + 1) + p(i, j - 1) \quad (14)$$

3.2. Averaging over rows

The signal $s_p(i,j)$ is a two-dimensional signal made from the second derivative signals for each row of the image. The magnitudes of the rows of second derivative signals are averaged together to form a pseudo-variance signal $v_p(j)$:

$$v_p(j) = \sum_{i=0}^R |d_p(i, j)| \quad (15)$$

This step is computationally faster than computing the variance signal (8) but yields similar results. It can be shown that the expected value of the absolute value of a zero mean normal distribution is $\frac{2\sigma}{\sqrt{2\pi}}$.

Figure 6 shows the mean second derivative signal for an original image and two interpolated images.

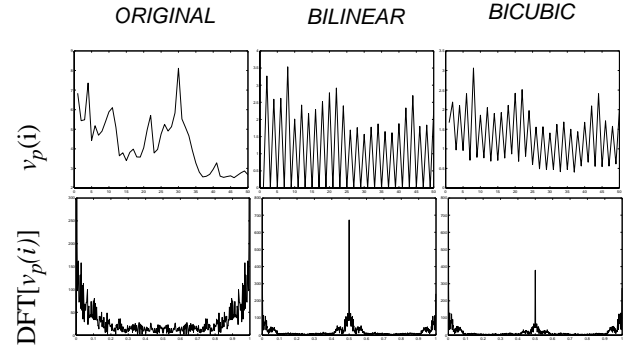


Figure 6: TOP ROW: Samples of the $v_p(j)$ signal from corresponding image portions for the cases of: a high-resolution original, bilinear interpolation by a factor of 2, bicubic interpolation by a factor of 2. **BOTTOM ROW:** The corresponding magnitude DFTs of the $v_p(j)$ signals. The high resolution original case does not have a signature DFT, the bilinear and bicubic interpolation cases each have a pronounced spike in the DFT. Note that the spikes occur at normalized frequency of 1/2, with a corresponding period of 2 (the estimated interpolation factor $\hat{N} = 2$).

3.3. Computing the DFT

When the image has been interpolated, the pseudo-variance signal exhibits a periodicity. This periodicity can be extracted via analysis in the frequency domain.

The DFT of a time or space signal is a frequency representation of the signal. Periodicity in the $v_p(i)$ signal may be determined by detecting peaks in the DFT[$v_p(i)$] signal. The DFT is computed with $C - 2$ points. The DFT of the signal $v_p(i)$ contains the frequency information at $C - 2$ points, representing normalized frequencies from 0 to $1 - \frac{1}{C-2}$, with an increment of $\frac{1}{C-2}$. Each frequency f (in units of cycles/pixel) has an associated period (in units of pixels/cycle) determined by the inverse of the frequency. A peak at any particular frequency in the DFT corresponds to a possible interpolation factor of $\hat{N} = \frac{1}{f}$.

Figure 6 shows the DFT signals corresponding to the $v_p(j)$ examples. In addition, Figure 7 shows several DFT signals corresponding to several different interpolation factors N for both bilinear and bicubic image interpolation. As in Figure 6, the bilinear interpolation generates a stronger peak than bicubic interpolation in the magnitude DFT of the $v_p(j)$ signal. The characteristic of the DFT is different, depending on whether bilinear or bicubic interpolation is used. For bilinear interpolation, spikes appear at each integer multiple of $1/f$ until $1/f > 1/2$. For bicubic interpola-

tion, a single spike at $1/f$ appears in the magnitude of the DFT.

3.4. Determining the interpolation factor

The interpolation factor determination stage of the algorithm consists of finding a peak f_p in the DFT and determining an estimated interpolation factor \hat{N} from f_p . Constants and thresholds were empirically determined, based on the nature of the DFT signal. The peak detection of the DFT includes the following steps:

1. Only a portion of the DFT spectrum is searched for peaks. Presently, the lowest frequencies of the spectrum are ignored (effectively limiting the estimated interpolation factor to a maximum of $\hat{N} = 9$). Because of symmetry, the spectrum is only searched for peaks up to normalized frequency of 0.5. Also, the peak detection operates exclusively on the magnitude of the DFT, ignoring phase.

2. A frequency is classified as a candidate peak if its associated magnitude is a local maximum and is T times greater than a local average magnitude. T is selected to be 10.

If the peak detection search yields no peaks, the algorithm concludes that no interpolation has been performed on the image. Otherwise, the peak f_p is the candidate peak having the greatest magnitude. From this peak, the interpo-

lation rate \hat{N} is estimated. The relationship between the peak f_p and \hat{N} is:

$$\hat{N} = \frac{1}{f_p} \text{ or } \hat{N} = \frac{1}{1-f_p} \quad (16)$$

Because of aliasing as described in Section 2, each peak corresponds to one of multiple interpolation factors. (For example, the peaks associated with an interpolation by 1.5 will appear similar to those of an interpolation by 3.0.) Because each point in the DFT may be interpreted as an interpolation factor greater than or less than 2.0, assumptions or a priori knowledge about possible values of N may solve the ambiguity. Perhaps, in some cases, additional clues from the digital image or its frequency representation could be used to determine the most probable of the possible rates of interpolation associated with f_p .

The number of points in the DFT determines the resolution with which the interpolation rate \hat{N} may be estimated. Because the estimated interpolation rate is derived from a frequency f in the DFT, by the relationship $\hat{N} = \frac{1}{f}$ and f is uniformly sampled, the estimated interpolation rate has finest resolution for low interpolation rates.

4. Detection of interpolation prior to JPEG compression

Digital images are typically compressed with JPEG. Often, interpolation of an image occurs before JPEG compression. The compression acts as a noise source, making the detection of interpolation more difficult. In some ways, JPEG decoding is similar to an interpolation algorithm. A JPEG compressed image contains one DC level for each 8×8 pixel block of the image. In the decoding process, this DC level is used (in conjunction with the coefficients of the DCT) to reconstruct the image. In the absence of the non-DC coefficients of the DCT, JPEG decoding has an effect similar to nearest neighbor interpolation. Therefore, it should be expected that any method used to detect an interpolation may also detect a factor of 8 interpolation when JPEG compression is present. Figure 8 shows the magnitude of the DFT $[v_p(i)]$ signal for a non-interpolated image compressed with JPEG. Peaks occur at normalized frequencies of $1/8, 1/4, 3/8, 5/8, 3/4$ and $7/8$. When analyzing JPEG compressed images with the interpolation detection algorithm, these peaks must be ignored. As a result, the peak associated with the true interpolation factor can be

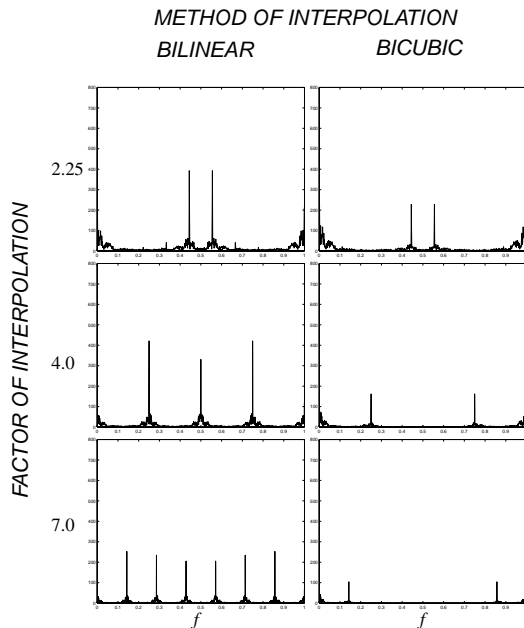


Figure 7: The magnitude DFT signals for a variety of interpolation factors and methods. Each row is a different interpolation factor (2.25, 4, or 7). Each column is a method of interpolation (from left to right: bilinear, bicubic).

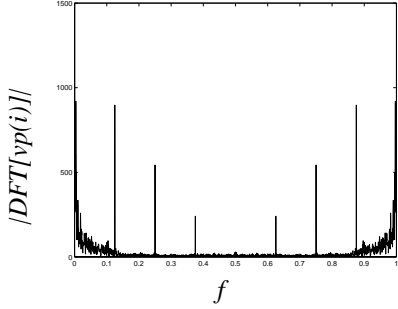


Figure 8: The magnitude DFT signal for a non-interpolated image compressed with JPEG compression.

found, but detecting interpolations by factors of 8, 4, 8/3, 8/5, 4/3, or 8/7 becomes more difficult.

Figure 9 shows the magnitude of the $DFT[v_p(i)]$ signal for an image created by an interpolation of 2.8 followed by JPEG compression. When the peaks associated with JPEG compression are ignored, the peaks at a normalized frequency of 0.357 are found. According to (16), this peak corresponds to an estimated interpolation factor of either 2.8 or 14/9.

5. Experiment

Instead of an optical zoom lens, the Kodak EasyShare CX7300 digital camera has a “digital zoom” feature that allows the user to zoom in on a region of the scene. Digital zoom is a feature commonly found in consumer digital cameras that essentially performs an interpolation of the image prior to image compression. In the CX7300, any zoom factor (i.e. interpolation factor N) from 1.1 to 3.0 in increments of 0.1 can be selected.

The zoom factor for a particular image can be found in the EXIF header under the tag “CaptureConditions.DigitalZoomRatio”

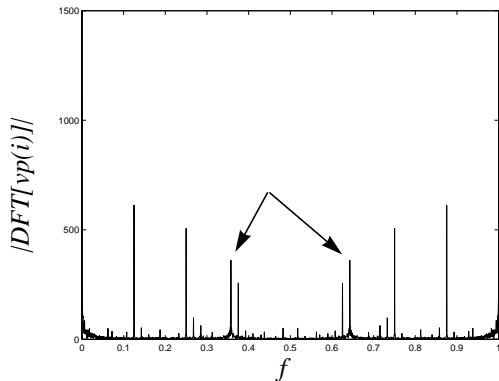


Figure 9: The magnitude DFT signal for an image compressed with JPEG following interpolation by a factor of 2.8. The arrows show the peaks associated with interpolation.

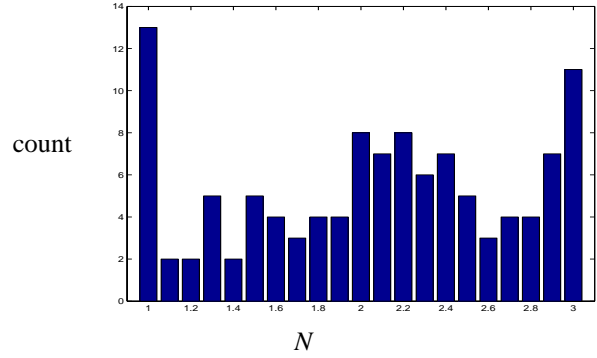


Figure 10: Histogram of the zoom factors for 114 images from the Kodak EasyShare CX7300 digital camera. A digital zoom value of 1.0 indicates that the image was not interpolated.

The interpolation detection algorithm was used to estimate the “digital zoom” amount used in the Kodak CX7300 digital camera.

The CX7300 was used to capture 114 images. The camera was set to capture at the highest resolution (1544×2080 pixels) and the highest quality JPEG compression. A histogram of the distribution of the digital zoom values is shown in Figure 10. Thirteen images did not use the digital zoom feature and were not interpolated.

As described in Section 3.4, as a consequence of aliasing, the estimate of the interpolation rate is one of two values (16). Because a priori knowledge of the possible interpolation factors exists, in most cases, the aliasing problem can be resolved. When the peak f_p is found, the two candidate estimated interpolation rates are determined. These candidate interpolation rates are examined against the list of possible interpolation factors. Often this step removes the ambiguity caused by aliasing by leaving only a single interpolation rate. For example, consider again the interpolated image having a DFT signal as shown in Figure 9. The candidate interpolation rates are 2.8 or 14/9. However, because it is known that the CX7300’s options for digital zoom rate are from 1.1 to 3.0 in increments of 0.1, we can eliminate 14/9 and estimate correctly that the image was interpolated by a factor of 2.8. For the CX7300, with one exception, at least one candidate interpolation rate can be eliminated based on the a priori knowledge. The exception is for interpolation factors of 1.5 and 3.0. When the true interpolation rate is either 1.5 or 3.0, the candidate interpolation rates are 1.5 and 3.0 and neither can be eliminated from consideration. In that situation, the ambiguity caused by aliasing cannot be resolved.

The interpolation detection algorithm was applied to all 114 images from the CX7300 and correctly identified all 13 images that were not interpolated. The interpolation factors for the 16 images interpolated by factors of 1.5 or 3

WINDOWED SINC INTERPOLATION BY 4.0

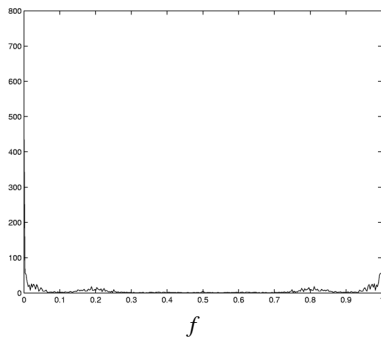


Figure 11: The $DFT[v_p(f)]$ signal associated with an interpolation factor of $N = 4$ for a windowed sinc interpolation. No peaks are noticeable. The interpolation detection algorithm performs poorly for high order interpolators.

were correctly, though ambiguously, identified as either 1.5 or 3.0. Finally, the interpolation factor estimate was correct for each of the remaining 85 images.

6. Algorithm limitations

The interpolation detection generally performs well when detecting interpolation by a low-order interpolation function (such as the linear or cubic interpolation function.) The algorithm works by detecting patterns in pixel-wise differences. These differences become more difficult to detect as each interpolated pixel value is determined as a function of more pixels in the original image. For example, for a particular windowed sinc interpolation function, each interpolated pixel value is a function of 225 pixels (15×15 pixel window) from the original image. For comparison with bicubic interpolation, each interpolated pixel value is a function of 16 original pixel values. The performance of the interpolation detection algorithm decreases as the order of the interpolator increases. To illustrate, an image was interpolated by a factor of 4 with the aforementioned windowed sinc interpolator. The resulting DFT signal with no detectable peaks is shown in Figure 11.

A second limitation of the interpolation detection algorithm is for a special case of interpolation by a factor of 2.0 where phase is preserved, an example is shown in Figure 12. Original sample locations (dashed lines) fall midway between interpolated sample locations (dotted lines). In this specific scenario, it can be shown that the sampled variance signal $v(x_0)$ is identical at each interpolated sample location. Therefore, the DFT signal will fail to produce meaningful peaks. Figure 13 illustrates the DFT signal for an interpolated image produced using a phase-preserving bicubic interpolation by a factor of 2.0. However, other

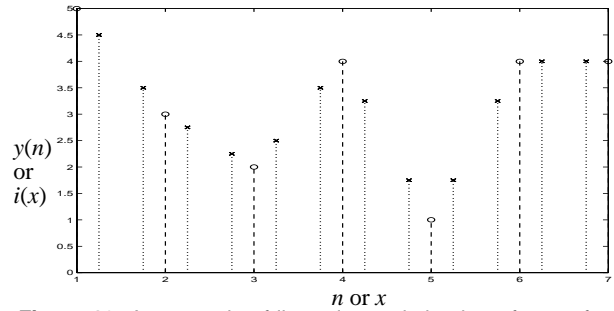


Figure 12: An example of linear interpolation by a factor of 2 while preserving phase.

phase-preserving factors of interpolation other than 2.0 are robustly detected by the algorithm, as described by Figure 5.

7. Conclusions

The novel interpolation detection algorithm is a fast and efficient algorithm for determining if an image has undergone interpolation with a low-order interpolator and the rate of that interpolation. The algorithm operates by exploiting the property that the second derivative signal of the interpolated images contains a periodicity. The algorithm produced reliable results on test images from a Kodak EasyShare CX7300 digital camera where interpolation occurs prior to compression with the digital zoom feature. The performance of the algorithm degrades for high-order interpolation filters such as a windowed sinc interpolation filter.

PHASE-PRESERVING BILINEAR INTERPOLATION
FACTOR 2.0

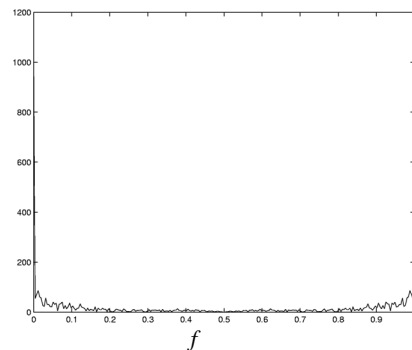


Figure 13: The $DFT[v_p(f)]$ signal associated with an interpolation factor of $N = 2$ for a bilinear interpolation with phase preservation. No peaks are noticeable, and the interpolation detection algorithm fails to detect the interpolation for the specific case of interpolating by a factor of 2 and preserving phase.

Overall, the algorithm has demonstrated robust results, and might it prove to be useful for situations where knowledge of an image's history determines the action of an image processing chain.

Kodak and EasyShare are trademarks of Eastman Kodak Company.

References

- [1] J. Proakis, D. Manolakis, "Digital Signal Processing: Principles, Algorithms, and Applications," MacMillan, NY, 1992.
- [2] R. G. Keys, "Cubic Convolution Interpolation for Digital Image Processing," IEEE Trans. Acoustics, Speech, Signal Proc, Vol. ASSP-29, No. 6, 1981, pp. 1153-60.
- [3] S. Etz, J. Luo, "Ground Truth for Training and Evaluation of Automatic 'Main Subject' Detection," SPIE Electronic Imaging Conference Proceedings, January 2000.