Systematic errors in digital image correlation caused by intensity interpolation

Hubert W. Schreier
Joachim R. Braasch
Michael A. Sutton
University of South Carolina
Department of Mechanical Engineering
300 Main Street
Columbia, South Carolina 29208
E-mail: sutton@sc.edu

Abstract. Recently, digital image correlation as a tool for surface deformation measurements has found widespread use and acceptance in the field of experimental mechanics. The method is known to reconstruct displacements with subpixel accuracy that depends on various factors such as image quality, noise, and the correlation algorithm chosen. However, the systematic errors of the method have not been studied in detail. We address the systematic errors of the iterative spatial domain cross-correlation algorithm caused by gray-value interpolation. We investigate the position-dependent bias in a numerical study and show that it can lead to apparent strains of the order of 40% of the actual strain level. Furthermore, we present methods to reduce this bias to acceptable levels.

Subject terms: image correlation; bias; systematic error; phase error; interpolation.

Paper 990442 received Nov. 8, 1999; revised manuscript received June 2, 2000; accepted for publication June 7, 2000.

1 Introduction

It has been observed in many of our experiments that the iterative cross-correlation algorithm\(^1\) introduces a systematic error depending on the subpixel position of the displaced subset. This error has been observed in pure translation tests as well as in strained specimens. For a long time, the errors remained unexplained and only recently has this error been attributed to the interpolation method used in the cross-correlation algorithm.\(^2\) A plot illustrating the typical relationship between systematic displacement error and subpixel position of the displaced subset is shown in Fig. 1. We can see that the error is zero at integer and midpoint positions. For a specimen uniformly strained in the \(\text{x}\) direction, where \(u(x) = e x,\) the relative apparent strain caused by this bias is found as

\[
\Delta e \equiv \frac{\Delta u}{\Delta (e x)} = \frac{\Delta u}{e}\Delta x,
\]

where \(\Delta u\) corresponds to the ordinate shown in Fig. 1 and \(e x = u\) corresponds to the abscissa. Thus, the relative apparent strain is given by the slope of the error function, and for the case shown, a maximum relative apparent strain of approximately 20% of the actual strain level can be expected, limiting the usefulness of cross-correlation for strain measurements.

2 Interpolation Errors

The cause of the position-dependent bias is found in the properties of the interpolators used in the iterative cross-correlation algorithm. To obtain subpixel accuracy, the cross-correlation coefficient (or other matching criterion, e.g., squared gray-value differences) must be evaluated at noninteger locations. Therefore, gray values and gray-value derivatives must be interpolated. It is well known that the most commonly used interpolators, namely, polynomial and B-spline interpolators, alter both the signal amplitude and the signal phase, depending on the subpixel position and the wave number of the signal.\(^3\) For a linear, shift-invariant interpolation filter with a transfer function \(\hat{h}(\hat{k}, \delta)\), the phase error \(\Delta \phi\) can be calculated by subtracting the linear phase shift \(\delta \pi \hat{k}\) caused by a positional shift \(\delta\) according to the shift theorem:

\[
\Delta \phi = \arctan \left( \frac{\text{Im} \hat{h}(\hat{k}, \delta)}{\text{Re} \hat{h}(\hat{k}, \delta)} \right) - \delta \pi \hat{k},
\]

where \(\hat{k}\) denotes normalized wave numbers. The positional error \(\Delta p\) corresponding to the phase shift \(\Delta \phi\) is given by

\[
\Delta p = \lambda \Delta \phi / 2\pi = \Delta \phi / \pi \hat{k}.
\]

As an example, a cubic polynomial interpolator can be expressed as a convolution operator with the kernel

\[
[h_0, h_1, h_2, h_3],
\]

where
Using Eqs. 3 and 4, the positional error as a function of subpixel position and wave number is found as

$$h_d = \frac{1}{2}\delta^2 - \frac{1}{16} + \frac{3}{2\pi} \delta - \frac{1}{8}\delta^3,$$

$$h_1 = -\frac{1}{4}\delta^2 + \frac{9}{16} - \frac{1}{2}\delta + \frac{1}{8}\delta^3,$$

$$h_2 = -\frac{1}{4}\delta^2 + \frac{1}{16} + \frac{1}{2}\delta - \frac{1}{8}\delta^3,$$

$$h_3 = \frac{1}{2}\delta^2 - \frac{1}{16} + \frac{1}{2}\delta + \frac{1}{8}\delta^3.$$  \hspace{1cm} (6)

Here, the polynomials are grouped into the even and odd parts and \(\delta\) denotes the subpixel position measured from the midpoint between samples for symmetry reasons. Using the symmetry of the coefficients, the corresponding transfer function is easily found as

$$\hat{h}_d(k, \delta) = (-\frac{1}{2}\delta^2 + \frac{1}{2})\cos(\frac{1}{2}\pi k) + (\frac{1}{4}\delta^2 - \frac{1}{2})\cos(\frac{3}{2}\pi k) + i[(\frac{9}{2}\delta - \frac{1}{2})\sin(\frac{1}{2}\pi k) - (\frac{1}{4}\delta - \frac{1}{4}\delta^3)\sin(\frac{3}{2}\pi k)].$$  \hspace{1cm} (7)

Using Eqs. (3) and (4), the positional error as a function of subpixel position and wave number is found as

$$\Delta u(k, \delta) = \frac{1}{\pi k} \arctan \left[ \frac{(9/4\delta - \delta^3)\sin(1/2\pi k) - (1/12\delta - 1/3\delta^3)\sin(3/2\pi k)}{(-1/2\delta^2 + 9/8)\cos(1/2\pi k) + (1/2\delta^2 - 1/8)\cos(3/2\pi k)} \right] - \delta. $$  \hspace{1cm} (8)

The positional error \(\Delta u\) and the amplitude attenuation \(|\hat{h}_d|\) are shown in Fig. 2 as a function of the fractional position \(\delta\) and the wave number \(k\). At the integer positions \(\delta=1/2\) and \(\delta=-1/2\), both errors vanish. At the midpoint, \(\delta=0\), the positional error is zero for symmetry reasons, but the amplitude error is highest. Note that the general shape of the error shown in Fig. 2 is the same for all polynomial and B-spline interpolators, only the magnitude of the errors changes.

The effect of the positional interpolation error on subpixel reconstruction is difficult to predict, as the error varies with frequency. For the case of reconstructing the displacement between two monochromatic waves, however, the effect can easily be imagined. The cross-correlation matching criterion is insensitive to a scale, such that the amplitude attenuation of interpolation has no effect. The maximum correlation will occur when the original wave and the wave interpolated from the displaced copy are in phase. This will be the case if the measured displacement \(u^*\) plus the positional error \(\Delta_p(u^*)\) is equal to the true displacement \(u_T\). Therefore, the reconstruction error \(\Delta u = u^* - u_T\) becomes

$$\Delta u = -\Delta_p(u^*).$$  \hspace{1cm} (9)

For a first-order approximation of the error as a function of the true displacement \(u_T\), one can assume that \(\Delta_p(u^*) \approx \Delta_p(u_T)\), and the reconstruction error becomes the negative of the interpolation error.

### 3 Numerical Studies

As the systematic errors of the cross-correlation results are influenced by many factors such as the frequency content of the speckle pattern, the amplitude attenuation and the phase error of the interpolator used, etc., an analytical pre-
diction of the errors cannot be found easily. Therefore, numerical studies with different speckle patterns and interpolation methods were performed. The correlation algorithm used is based on an iterative search algorithm to find the maximum of the cross-correlation coefficient $r$. For simplicity, suppose the origin of an $x,y$ coordinate system is located at the center pixel of an $(N+1) \times (N+1)$ subset in the first image $f(x,y)$, where $N$ is even. To find the corresponding subset in the second image $g(x,y)$, the cross-correlation coefficient

$$r = \frac{\sum_{y=-N/2}^{N/2} \sum_{x=-N/2}^{N/2} f(x,y) g(\xi(x,y,s), \eta(x,y,s))}{\sqrt{\left(\sum_{y=-N/2}^{N/2} \sum_{x=-N/2}^{N/2} f^2(x,y)\right) \left(\sum_{y=-N/2}^{N/2} \sum_{x=-N/2}^{N/2} g^2(\xi(x,y,s), \eta(x,y,s))\right)}}$$

(10)

is defined. In this equation, $s$ is a parameter vector that relates coordinates in the first image subset to the corresponding coordinates in the second image through

$$\xi(x,y,s) = s_1 x + s_2 y + s_3 y,$$

(11)

$$\eta(x,y,s) = s_4 x + s_5 y + s_6 y.$$  

(12)

The parameters $s_1$ and $s_4$ describe the $u$ and $v$ displacement components between the subset centers, respectively, and the remaining parameters allow for an affine transform. To find the parameter vector $s_{\text{max}}$ that maximizes Eq. (10), the cross-correlation coefficient is developed into a second-order Taylor polynomial at a point $\bar{s}$ in the vicinity of the correlation peak. Assuming the Taylor polynomial is correct, the maximum can be found by

$$s_{\text{max}} = \bar{s} + \Delta \bar{s},$$  

(13)

where $\Delta \bar{s}$ is obtained from

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & r_{15} & r_{16} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \Delta s_1 \\ \Delta s_2 \\ \Delta s_3 \\ \Delta s_4 \\ \Delta s_5 \\ \Delta s_6 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \end{bmatrix},$$  

(14)

Here, $r_i$ denotes $\partial r / \partial s_i |_{\bar{s}}$ and $r_{ij}$ denotes $\partial^2 r / \partial s_i \partial s_j |_{\bar{s}}$. Equations (13) and (14) are used to iteratively find $s_{\text{max}}$ in a Levenberg-Marquardt algorithm. The program "Vic2D" was used for all numerical studies.

The speckle images used for the numerical studies were taken from a number of previous experiments. To isolate the effects of interpolation, translated and stretched images were generated numerically from the speckle images. Of course, the generated images must not have an interpolation bias. Therefore, the images were obtained by applying the appropriate shift in the Fourier domain according to the shift theorem. For a given subpixel displacement $\Delta x$, an interpolated value can be calculated by shifting the signal by $-\Delta x$, which brings the sought value to an integer loca-

4 Results

Initial studies with linear interpolation indicated that relative apparent strain errors in excess of 40% can be introduced by the phase errors. This is not surprising, as linear interpolation causes a maximum positional error of approximately 1/20 of a pixel for a periodic structure of 4 pixels wavelength, and 0.13 pixels error for a wavelength of 3 pixels. As the magnitude of these errors is clearly unacceptable, all further investigations were limited to cubic polynomial, cubic B-spline and quintic B-spline interpolation.

Extensive studies were performed for a large number of different speckle patterns. It was consistently found that speckle patterns with a bimodal gray-value distribution produced the highest errors, while patterns with a uniform distribution exhibited significantly less error. This can be readily attributed to the higher interpolation errors toward high wave numbers. While the gray-value distribution is not an exact measure of frequency content, it is obvious that a bimodal distribution with essentially black speckles on a white background (or vice versa) has more energy in the high wave number range than an image with smooth transitions between black and white. To limit the amount of data presented, we chose two speckle patterns. The speckle pattern shown in Fig. 3(a) had the least amount of error of all patterns investigated. This image approximates a continuously varying intensity pattern and has a uniform gray-level distribution. The second image chosen was produced by mapping the lower half of the gray values to the lower quarter and the upper half to the upper quarter, thus producing a bimodal gray-value distribution [see Fig. 3(b)]. The results from the image generated in this manner were chosen because they can be regarded as an upper bound of the errors that occurred with images taken from actual experiments. Furthermore, the characteristics of both images, particularly speckle size and orientation, are comparable.

4.1 Translation

First, a series of translated speckle images was investigated. The shifted images were computed by applying a Fourier filter $\exp(-i n \pi \Delta x)$, where $n$ indicates the image number and $\Delta x$ is the shift increment between successive images. In this manner, a series of 20 images was generated, corresponding to a shift of 0.05 pixel between successive images. The images were then analyzed with the iterative
cross-correlation algorithm described in the previous section. The displacements were calculated at one thousand points for every image with a subset size of 31×31 pixels. Figure 4 shows the correlation error as a function of the true subpixel displacement for the image with the continuous gray-value distribution on the left and for the bimodal image on the right. While quintic spline interpolation produces the least amount of error for the continuous image, as one would intuitively assume, the errors are larger than for the other interpolators for the bimodal gray-value distribution. These somewhat counterintuitive results clearly point out the importance of both phase error and amplitude attenuation for the matching process. Even though the quintic spline interpolator has smaller phase errors for a given wave number, it can produce worse results because it has less amplitude damping and does not filter out the high wave numbers to the same degree as lower order interpolation. The cubic polynomial interpolation exhibits the lowest error for the bimodal image. The iterative minimization process becomes slightly unstable for the bimodal case, with the number of iterations being approximately three times higher than for the continuous image. In this case, higher random errors occur, which cancel due to averaging, rather than the pronounced systematic error found for the continuous gray-value distribution. The instability can be attributed to errors in the second order derivatives of the cross-correlation coefficient, which are approximated using only the first-order gray-value derivatives for efficiency reasons.

4.2 Strain

As a second test, strained images were generated to investigate how the systematic errors are affected by a nonuniform displacement field. The generated images had a uniform strain of 0.005 in the horizontal direction of the image. Twenty-one columns of the image, corresponding to subset displacements between 0 and 1 pixels in 0.05-pixel increments, were analyzed with the correlation program. As the true displacement was constant along vertical lines in the image, the average was taken along the vertical direction, where 450 data points were available for each subpixel displacement, resulting in a total of 9450 points analyzed. Figure 5 shows the average error in the reconstructed displacement over the true subpixel displacement. The errors are on the same order as for uniform displacement for both the continuous and the bimodal gray-value distribution. The same instability noted previously occurred for the cubic polynomial interpolation and the bimodal pattern.

As mentioned in the introduction, the relative apparent strain error caused by the bias is given by the slope of the error plots shown in Fig. 5. Figure 6 shows the relative apparent strains obtained by numerical differentiation of the
the displacement errors. For the continuous pattern, errors of the order of 4% occur for cubic polynomial interpolation, and for the bimodal pattern, quintic spline interpolation produces errors of the order of 15%.

Similar tests were performed for different strain levels ranging from 0.02 to 5%. The errors do not show any dramatic change with strain level; only toward the higher end was a slight decrease in error observed.

4.3 Error Reduction
As the interpolation errors increase with the wave number \( \tilde{k} \), the reduction of the high-frequency content in the speckle images by low-pass filters can be used to reduce systematic errors. At first, this procedure appears to reduce the information content that is desirable for correlation. One would expect the cross-correlation algorithm to rely on the presence of pronounced edges between black and white speckles, which are blurred by a low-pass filter. However, due to the nature of common speckle patterns, the information content in the low-frequency range is very high and sufficient for accurate matching. Applying a low-pass filter can be thought of as a weighting function that reduces the erroneous contribution of the high wave numbers to the correlation result. The spatial resolution is unaffected by this procedure, as long as the filter cutoff is sufficiently separated from the frequency band of the displacement functions. Figure 7 shows the correlation error for the translated and strained bimodal speckle patterns for a binomial [0.25, 0.5, 0.25] low-pass filter applied prior to correlation. The error reduction for spline interpolation is significant, and quintic spline interpolation shows approximately five times less error than cubic spline interpolation. Since cubic interpolation is stable in this case, the algorithm converges to the biased location and the systematic errors are pronounced.

For the plot on the right-hand side in Fig. 7, which was obtained from the strained bimodal speckle pattern, the relative apparent strain error is given by the slope of the error curves shown. Thus, the error reduction is not limited to displacement errors, but also provides a substantial reduction of the strain errors. For quintic spline interpolation, the relative apparent strain error obtained after numerical differentiation of the displacement errors was less than 2% for all cases studied.

4.4 Convergence and Speed
Cross-correlation algorithms that perform the search for the correlation peak in the spatial domain are commonly regarded as extremely slow compared to Fourier domain methods.\(^6\) This stems largely from early implementations

![Fig. 6](http://opticalengineering.spiedigitallibrary.org/) Relative apparent strain for different interpolation functions: left, continuous gray-value distribution; right, bimodal distribution.

![Fig. 7](http://opticalengineering.spiedigitallibrary.org/) Systematic error for the bimodal pattern using binomial low-pass filters: left, pure translation; right, uniaxial strain.
using a coarse-to-fine search strategy,\textsuperscript{7} which has since been replaced by Newton-Raphson or Levenberg-Marquardt methods.\textsuperscript{5} Although Fourier processing is faster than the older spatial coarse-fine search methods, note that Fourier methods exhibit a similar systematic error as the spatial domain methods relying on interpolation.\textsuperscript{9} This bias is caused by peak estimators used to obtain a subpixel estimate of the peak location from the cross-correlation function sampled at integer pixel locations.

Using the Levenberg-Marquardt search method, the speed of spatial domain algorithms has become, at the very least, comparable to the Fourier methods. This is due to the reduction of points where the cross-correlation function has to be evaluated. Typically, the average number of iterations varies between two and five, depending on factors such as noise, strain, rotation, and very importantly, the interpolation method used. At first glance, linear interpolation appears to be the fastest choice, as only four multiplications are required to interpolate a gray value, compared to 20 multiplications for a cubic scheme. However, the search algorithm converges much faster for cubic than for linear interpolation, which, in most cases, compensates for the additional multiplications used for interpolation. In addition, calculating the derivatives of the cross-correlation coefficient is more time-consuming than the gray-value interpolation and differentiation while being totally independent of the interpolation method used. Even though quintic interpolation requires more than twice as many multiplications as cubic interpolation, the speed reduction is only about 20\% on average, with the number of iterations being approximately the same. As the number of points where displacements are evaluated is typically of the order of at least several hundred points, the additional computation time required for the spline transform is negligible, as it can be very efficiently implemented due to the work of Unser et al.\textsuperscript{10, 11} The same argument applies to the additional computational cost of a binomial low-pass filter.

In summary, cubic and quintic spline interpolation are preferable to polynomial interpolation because of the better convergence and smaller phase errors. This benefit is achieved at virtually no additional computational cost. Only if very few points are evaluated and the speed requirement is critical, e.g., in an on-line inspection system, can cubic polynomial interpolation be beneficial.

5 Discussion

To further illustrate the impact of the systematic error and to obtain insight into the potential of the method for small strain determination, we investigated the following example. According to the method described above, a uniformly strained image with an applied strain of 50 \( \mu \varepsilon \) was created from the test image with the continuous gray-value distribution. One line of displacement data spanning 450 columns of the image was analyzed with different interpolation methods. The strain was then calculated from the displacement data as the slope of the best linear fit. The results were 59.8, 53.4, 52.8, and 50.1 \( \mu \varepsilon \) for linear, cubic polynomial, cubic B-spline, and quintic B-spline interpolation, respectively.

These results suggest that the systematic errors due to intensity interpolation can be virtually eliminated. Furthermore, they show that digital image correlation methods have no inherent systematic errors preventing them from making strain measurements with accuracies comparable to those of strain gages, provided that appropriate signal processing techniques are employed. However, there are other sources of error that affect the accuracy of the method that have not been taken into consideration by this study. Aside from experimental factors, the method is currently somewhat limited by the random noise in the displacement data. In the preceding example, the strain exhibits random variations of \( \pm 10 \) and \( \pm 50 \mu \varepsilon \) if gage lengths of 100 and 31 pixels, respectively, are used for numerically differentiating the displacement data obtained by quintic B-spline interpolation. To mitigate the effect of random noise on the strain measurements, we are currently investigating a method of incorporating smoothness constraints into the correlation algorithm directly, rather than relying solely on post-processing of the displacement data.

The impact of the interpolation error is not limited to small strain measurements. The absolute strain errors due to interpolation scale linearly with the strain level. Therefore, attention to the interpolation error should be paid in all cases. We feel that the proposed method of error reduction by high-order interpolation functions and low-pass filtering comes at very little additional computational expense for the error reduction that can be achieved. Particularly, we see no justification for continued use of linear interpolation, as the subpixel information gained by this method is severely biased. If care is taken, the relative apparent strain introduced by cubic interpolation methods can be limited to below 5\%, and to below 2\% if quintic B-spline interpolation is used.

Finally, the relationship between this work and previous numerical simulations\textsuperscript{4, 5} was studied. Sutton et al. showed that there was a bias in the numerical simulation results that varied between being parabolic and sinusoidal on the interval \( u \in [0,1] \). Their simulation results included the effects of interpolation errors, partial sensor sensitivity and quantization errors. Since our work has focused on defining and mitigating the effects of interpolation error, it is difficult to directly relate the results. However, in one regard, both current and previous work are in agreement; higher order interpolation methods are effective in reducing systematic errors in digital image correlation.

6 Conclusion

We showed the impact of interpolation phase error on the matching error in digital image correlation. This systematic error can cause significant apparent strains in data obtained through image correlation methods.

To reduce the systematic errors, high-order interpolation methods with smaller phase errors are preferable. The most dramatic error reduction is found in going from linear to cubic interpolation, while quintic interpolation showed a less pronounced improvement over cubic interpolation.

We also showed the influence of frequency content in the speckle pattern on matching bias and have found a smooth transition between black and white to be preferable for accurate measurements. Alternatively, low-pass filtering the speckle images prior to correlation can be used to limit the effect of phase error at high wave numbers that is characteristic of all interpolators with compact support. As digital image correlation is most commonly used as an off-line
processing tool with no particular speed requirements, the use of the somewhat slower quintic spline interpolation in conjunction with a low-pass filter is highly recommended.

Acknowledgments

We would like to thank Prof. Dr. Herbert Weber and Dipl.-Phys. Bernd Gutmann at the University of Karlsruhe for their support. We would also like to thank the Correlated Solutions’ staff for modifying their software according to our requests. In addition, the financial support through National Aeronautics and Space Administration (NASA) EPSCoR grant No. NCC5-174 and National Science Foundation (NSF) EPSCoR grant No. NSF-EPS-9630167 is deeply appreciated.

References


Hubert W. Schreier received his Diploma in process engineering from the University of Karlsruhe, Germany, in 1997. His research work at the University of Karlsruhe focused on the design and implementation of optical measurement systems for the analysis of surface waves. In 1997, he joined the research group of Prof. Michael A. Sutton at the University of South Carolina as a PhD student. He is currently working on the development of an optical measurement system for three-dimensional deformation measurements on the micro-scale. Other interests include the modeling and optimization of white-light speckle measurement systems, fringe projection, camera calibration and image processing in general.

Joachim R. Braasch received his Diploma in process engineering (ME) from the University of Karlsruhe, Germany, in 1999. In 1998, he joined Prof. Michael A. Sutton’s research group for a half year visit to work on the optimization of the accuracy of digital image correlation algorithms. Previously he had stayed at the University of Manchester Institute of Science and Technology (UMIST), Great Britain. He had joined the research team of Dr. Cilliers at the department of chemical engineering to work on the implementation of digital image processing in the control of the flotation process in the mining industry. Currently, he is employed by the Babcock Borsig AG, Oberhausen, in Germany and working for its subsidiary company, the Babcock Borsig Power Energy GmbH.

Michael Albert Sutton received his PhD from the University of Illinois. He is a Carolina Distinguished Professor in mechanical engineering, director of the State Center for Mechanics, Materials and Non-Destructive Evaluation, NASA Space Grant Campus Director, and interim chair of mechanical engineering. He joined the Mechanical Engineering Department at the University of South Carolina in 1982. He has co-authored nearly 100 journal papers, over 100 conference articles and four book chapters in the areas of computer vision, fracture mechanics and optical methods. He was elected a fellow of the Society for Experimental Mechanics in 1999. He has received three national awards from the Society for Experimental Mechanics and was designated a Presidential Young Investigator by the National Science Foundation in 1985. He served as chairperson of mechanical engineering from 1997 to 1998 and will be president of the Society for Experimental Mechanics in 2001.