

Automated Fast Initial Guess in Digital Image Correlation

Z. Wang*, M. Vo[†], H. Kieu* and T. Pan[‡]

*Department of Mechanical Engineering, The Catholic University of America, Washington, DC 20064, USA *Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA *Department of Civil Engineering, The Catholic University of America, Washington, DC 20064, USA

ABSTRACT: A challenging task that has hampered the fully automatic processing of the digital image correlation (DIC) technique is the initial guess when large deformation and rotation are present. In this paper, a robust scheme combining the concepts of a scale-invariant feature transform (SIFT) algorithm and an improved random sample consensus (iRANSAC) algorithm is employed to conduct an automated fast initial guess for the DIC technique. The scale-invariant feature transform algorithm can detect a certain number of matching points from two images even though the corresponding deformation and rotation are large or the images have periodic and identical patterns. After removing the wrong matches with the improved random sample consensus algorithm, the three pairs of closest and non-collinear matching points serve for the purpose of initial guess calculation. The validity of the technique is demonstrated by both computer simulation and real experiment.

KEY WORDS: digital image correlation, improved random sample consensus, initial guess, scale-invariant feature transform

Introduction

Digital image correlation (DIC) technique is one of the most widely used methods for shape, motion and deformation measurements [1]. The DIC technique typically works by comparing and matching the grayscale images of an object captured from different views, at different time, or at different stages of deformation. Through tracking every few pixels of interest in the reference and target images, followed by carrying out data interpolation, the DIC technique can determine the whole-field two-dimensional (2D) and three-dimensional (3D) shape, deformation and motion vector fields as well as their gradient maps. Because of its primary advantages of easy implementation and wide range of measurement sensitivity and resolution, the DIC technique has found numerous applications in many fields [2].

The DIC technique generally employs a correlation criterion to detect the best image matching for a group of pixels (named subset) centred at each interrogated pixel. There exist a variety of correlation criteria for the DIC analysis. For instance, a simple yet robust criterion named the parametric sum of squared difference criterion can be written as [3]

$$C = \sum_{i=1}^{N} \left[a f(x_i, y_i) + b - g(x'_i, y'_i) \right]^2$$
(1)

where *a* is a scale factor, *b* is an offset of intensity and $f(x_i, y_i)$ and $g(x'_i, y'_i)$ indicate the intensity values at the *i*th pixel in the reference subset and the matching pixel in the target subset, respectively. The task of the correlation

analysis is to minimise the coefficient *C* in Equation (1) to find the best matching. For a representative pixel (x_0, y_0) to be analysed in the reference image, a square pattern of $N = (2M+1) \times (2M+1)$ pixels with its centre located at (x_0, y_0) is usually chosen as the reference subset. The corresponding subset in the target image, i.e., the target subset, is often of irregular shape. Denoting the shift amount between the centres of the two matching subset patterns as (ζ, η) , the shape mapping function for the entire reference and target subsets can be expressed as

$$\begin{aligned} \dot{x'_i} &= x_i + \xi + \xi_x(x_i - x_0) + \xi_y(y_i - y_0) \\ \dot{y'_i} &= y_i + \eta + \eta_x(x_i - x_0) + \eta_y(y_i - y_0) \end{aligned}$$
 (2)

where ξ_x , ξ_y , η_x and η_y are the coefficients of the shape function. To determine all the six unknowns of shape function (ξ , η , ξ_x , ξ_y , η_x , η_y) as well as the scale and offset parameters (*a*,*b*) involved in Equation (1), the DIC technique often employs an iterative algorithm such as the Newton–Raphson or the Levenberg–Marquardt method to carry out the correlation optimization. The iterative algorithm is capable of providing very fast and highly accurate correlation analysis [4]. The downside, however, is that a reasonably good initial guess for the six unknowns of the shape function on the starting or seed point is required.

When the shape change and the rotation of the target image with respect to the reference image are relatively small, the initial values of ζ_x , ζ_y , η_x and η_y can be set to zeros. In this case, the initial values of ζ and η can be automatically

detected by a full-field subset scanning process [5, 6]. On the contrary, when the shape change and/or the rotation are relatively large, the initial guess normally has to be conducted by manually selecting three pairs of matching pixels in the reference and target images through humancomputer interaction, which yields six equations to solve for the six unknowns in Equation (2) as an initial guess. Another notable case where a manual initial guess is usually demanded is when each of the reference and target images has multiple identical or nearly identical regions; the reason is that the existing automatic initial guess methods may not be able to detect the correct matches in those regions. It is also noted that the shape mapping function in Equation 2 may have more than six parameters to include higher-order terms; nevertheless, the higher-order terms are negligible for the initial guess purpose.

Manual selection of points for the initial guess in the DIC analysis has hampered the fully automatic analysis feature of the technique. The situation can become worse when there are multiple regions of interest, each requiring a reliable initial guess. In this paper, a robust scheme combining the concepts of a scale-invariant feature transform (SIFT) algorithm and an improved random sample consensus (iRANSAC) algorithm is employed to achieve an automated fast initial guess for the DIC technique. The SIFT algorithm is widely used in computer vision to detect and describe local features in images, and it has been recently applied to deformation measurements [7, 8]. Despite that, the advantages of the SIFT algorithm for the DIC measurements have not been fully utilised. The novel approach to be presented has the ability to accurately and automatically detect a number of matching points from two images even though the deformation and rotation are large or the images have periodic and identical patterns.

Principle

Finding correct correspondences between images captured from different directions is traditionally a difficult problem. Recently, important advances have been made to tackle this problem. Among a few of them, the SIFT feature is one of the most important [9]. The essence of the SIFT method is twofold: (i) identify feature locations in the image space that are invariant with respect to the image translation, scaling and rotation; (ii) compute local descriptor around each feature point in the image that captures saliency of the point. The SIFT feature allows the correspondences of feature points between two images to be easily detected. The best matching pairs can be further identified, and the wrong matches can be eliminated by the iRANSAC algorithm.

Scale-invariant feature transform feature points

Keypoints detector

To achieve scale invariance, the interest points are obtained from scale-space extrema of the difference of Gaussian (DoG). This can be conducted efficiently by constructing an image pyramid in a procedure as follows. First, the Gaussian pyramid is formed from the original image by successive smoothing and subsampling. Second, the DoG pyramid is built from the differences of the subsampled images and its Gaussian smoothing version at that pyramid level, as illustrated in Figure 1. Third, the extrema of the scale-space function are detected by comparing each pixel in the pyramid image with its 26 neighbour pixels. Last, a quadratic polynomial is fitted to the local sample points of each detected extremum to localise it with a higher accuracy. This refinement is particularly important to estimate more accurately the scale information that is used for scale normalisation to achieve a scale invariant feature. The interpolation fitting function is also useful in determining the stability of the local extremum.

Keypoints descriptor

The SIFT descriptor exploits the local gradient information of image intensities to summarise the image appearance in a local neighbourhood around each interest point. To obtain rotational invariance, each interest point is assigned an orientation corresponding to the dominant direction computed from a local histogram of gradient directions accumulated over a neighbourhood of the point. It is noted that before being added to the histogram, each sample's gradient magnitude is weighted by a Gaussian window centred at the interrogated point with the size proportional



Figure 1: Scale-space pyramid. The image is subsampled by two in every octave

to its detection scale. Next, a 4×4 square grid of size proportional to the detection scale of the interest point centred at the interest point with its orientation determined by the dominant orientation is laid out in the image, as demonstrated in Figure 2. An eight-bin accumulated histogram is then generated from the image gradient directions within each grid. As a whole, these local histograms give rises to an image descriptor of $4 \times 4 \times 8 = 128$ dimensions for each interest point. This resulting image descriptor is referred to as the SIFT descriptor. In order to improve the robustness of the descriptor to illumination invariance, the 128-dimensional histogram is normalised to unit sum.

Feature point matching

With a set of image descriptors, the correspondences between images can be found by mutually matching their descriptors and choosing the one that minimises the Euclidean distance between them. While this procedure is non-deterministic polynomial-time hard, it is speeded up by using the fast library for approximate nearest neighbours (FLANN) with minimal loss of accuracy [10]. To suppress possibly ambiguous matches, it is recommended to only accept matches for which the ratio between the distances to the nearest and the next nearest points is less than 0.8 [9].

Improved random sample consensus matching

Although the SIFT descriptor provides a convenient way to match the correspondences between images, it is almost inevitable that mismatches occur if the appearance of the images change drastically or repetitive patterns are observed in each image. One solution to this issue is to apply the RANSAC algorithm [11] with some predefined model functions to remove those mismatches. This can be illustrated by the line fitting shown in Figure 3. Over the time, the original RANSAC algorithm has been remarkably improved in a variety of ways. Two of these improvements can be referred to as progressive sample consensus (PROSAC)



Figure 2: Scale-invariant feature transform descriptor. This descriptor contains 4×4 descriptor array calculated from 16×16 samples



Figure 3: Illustration of line fitting using random sample consensus to select reliable matchings

algorithm [12] and group sample consensus (GroupSAC) algorithm [13], which are briefly described below. The advantages of them are combined into an algorithm called iRANSAC in this paper.

PROSAC

Rather than blindly sampling candidate points to compute model fitting and checking for consistency of other points with respect to the computed model as done in the classical RANSAC algorithm, the PROSAC algorithm utilises the relative accuracy information among the data points to accomplish an efficient sampling scheme. Given the SIFT feature descriptors, the relative accuracy information can be readily obtained from the Euclidean distance between each descriptor. Although many features from an image will not be correctly matched, the tendency of similarity to better predict the correct correspondences than the random guessing usually holds true because shorter distance pairs have a significantly higher probability of being the correct matches. Based on this information, a growth function that progressively draws data from large sets of top-ranked potential candidates classified by the relative accuracy metric was introduced. In spite of not being analytically verified, this growth function had been empirically tested to be at least equally likely to find the optimal solutions as the RANSAC algorithm [12]. Effectively, this sampling scheme allows the more likely uncontaminated samples to be examined first, which leads to large computational savings over the traditional RANSAC algorithm. The termination criterion for the iterative procedure of the PROSAC algorithm is based on two constraints: (i) a maximality constraint determining when the probability of the existence of better solutions becomes lower than a predefined threshold; (ii) a non-randomness constraint ensuring that incorrect model associated with accidental outlier data will be discarded.

GroupSAC

The underlying assumption behind the GroupSAC algorithm is the existence of some correlations between correspondences

so that they can be clustered into groups. A simple way to group the potential matching candidates together is to apply optical flow on the offsets between the potential corresponding points in the images. The GroupSAC sampling paradigm is designed to take full advantage of the properties of clusters of data: data points within a group are consistent, and the group with more tentative matches will have a higher ratio. In addition, significant computation cost can be saved if the sampling scheme can find the best fit model from fewer groups. To this end, the novel sampling scheme was described as follows: (i) the groups are sorted in descending order of the number of data points; and (ii) the number of groups is progressively taken to draw samples from the union of these group configurations. The maximum number of iterations for each subset of the union can be computed by invoking the inclusion-exclusion principle in combinatorics on that subset. More progressive, GroupSAC goes through all possible configurations in the order of their cardinalities, i.e., number of data points in that group. As the number of groups chosen approaches to the total number of groups, each minimum sample set has equal chance of being sampled, and the performance of the GroupSAC algorithm gracefully degrades to that of the traditional RANSAC algorithm. The terminating conditions for the GroupSAC algorithm are the same as those required by the PROSAC algorithm.

Improved random sample consensus

Since the GroupSAC algorithm affects the sampling strategy only in a macro view, i.e., clustering the data into groups, it is natural to apply the PROSAC algorithm to each group configuration to efficiently compute the best fit model. The resulting iRANSAC algorithm comprises the following steps:

- 1. Cluster all the data into groups by using optical flow computed with respect to their offset between potential matches.
- 2. Sort these groups with respect to their numbers of data points.
- 3. Iterate until the number of iteration exceeds a predefined value or the convergence criterion is satisfied as follows:
 - (a) Progressively select the number of groups to form a group configuration.
 - (b) Data within the union of the above group configuration are sorted with respect to their relative accuracy information.
 - (c) Compute the maximum number of iteration needed for each subset of the configuration.
 - (d) Iterate until the number of iteration for each configuration subset exceeds the above computed value or the convergence criterion is satisfied as follows:

- i. Progressively sample data points from every subset of the configuration to compute the fitting model.
- ii. Check for consistency of all the data points with respect to the fitting model.
- iii. Terminate the iRANSAC process if the maximality and non-randomness constraints are satisfied.

Scale-invariant feature transform-based improved random sample consensus matching for digital image correlation initial guess

For the initial guess in the DIC applications, it usually suffices to model the transformation between the reference and target images as a homography transformation. Therefore, the fitting model can be of the following form:

$$\begin{cases} u_1 \\ v_1 \\ 1 \end{cases} \equiv H * \begin{cases} u_2 \\ v_2 \\ 1 \end{cases}$$
 (3)

where (u_1, v_1) and (u_2, v_2) are candidate points in the reference and target images, respectively. *H* is a 3×3 matrix that represents the coefficients of the homography



Figure 4: Flowchart of the digital image correlation (DIC) analysis with scale-invariant feature transform-improved random sample consensus (SIFT-iRANSAC)-based automated initial guess



Figure 5: Simulation analysis of images with large deformation and rotation (A) is the reference image and (B) is the target image. (C) and (D) are the detected matching pairs of points in the reference image and the target image, respectively; the adopted three matching pairs are surrounded by dotted ellipses. (E) and (F) are the DIC-detected motion of pixels in the horizontal and vertical directions, respectively; the plot is superposed on the initial reference image

transformation. Accordingly, the procedure of the proposed technique combining the SIFT and the iRANSAC algorithm for automated fast initial guess is comprised of four steps:

- 1. Extract the SIFT keypoints and descriptors for both images.
- 2. Run the FLANN algorithm to determine potential correspondences.
- 3. Run the iRANSAC algorithm to obtain correct matches from those potential correspondences. In practice, the number of eventual matching pairs can be set to no less than 30 so that there are sufficient points to select for the DIC initial guess.
- 4. In the reference image, three closest matching points in the region of interest that can form a regular-shaped

triangle are adopted, together with the matching points in the target image, to find the initial guess to start the DIC analysis. Any of the three points can be chosen as the starting or seed point.

In step 4, the reason that the three closest matching points are adopted is that it provides a simple criterion to select the points, and reliable initial guess can be acquired for the measurements involving large strain variations. If the deformation gradient or strain distributions are relatively uniform, higher accuracy of initial guess could be achieved by selecting three matching points far away from each other. However, such accuracy enhancement on initial guess is negligible in practice because the subsequent DIC analysis commonly employs the robust



Figure 6: Simulation analysis of images with periodic patterns. (A) is the reference image and (B) is the target image. (C) and (D) are the detected matching pairs of points in the reference image and the target image, respectively; the adopted three matching pairs are surrounded by dotted circles. (E) and (F) are the digital image correlation-detected motion of pixels in the horizontal and vertical directions, respectively

Newton–Raphson or Levenberg–Marquardt iterative algorithms. For the situation where the three points are very close, e.g., 1–2 pixels apart from each other, it is expected that the estimation error can be a little larger. According to the numerous tests that were performed during the study, this problem never brings a problem to the eventual analysis because of the robustness nature of the subsequent DIC iteration process.

Figure 4 illustrates the flowchart of the DIC analysis with SIFT-iRANSAC-based automated initial guess. It may be helpful to note that the SIFT, RANSAC and extended RANSAC (e.g. PROSAC and GroupSAC) algorithms are typically not easy to understand; meanwhile, detailed descriptions could lead to redundancy and inappropriateness. For this reason, the relevant algorithms are briefly described in this paper with the intent of clear explanation.

Simulation and Experiment

Computer simulation tests have been carried out to verify the validity and effectiveness of the proposed approach. In the first simulation, a speckle pattern of 512×512 pixels serves as the reference image, and it is deformed as the target image by using the following equation:

$$X = 150.25 - 0.25x - 0.35y$$

$$Y = -100.75 + 0.5x + 0.25y$$
(4)

where (x, y) and (X, Y) denote the pixels in the reference and



Figure 7: Analysis of experimental images with large translation and rotation. (A) is the reference image and (B) is the target image. (C) and (D) are the detected matching pairs of points in the reference image and the target image, respectively; the adopted three matching pairs are surrounded by dotted circles. (E) and (F) are the digital image correlation-detected motion of pixels in the horizontal and vertical directions, respectively

target images, respectively. From the undeformed reference and deformed target images shown in Figure 5(A and B), it is evident that large deformation and rotation have occurred in the deformed image. This conventionally requires a manual selection of three matching pairs of points to find the initial guess for the DIC analysis.

Figure 5(C and D) shows the 112 matching points detected by the presented SIFT and iRANSAC algorithms, and three pairs of points that are non-collinear and close to each other have been selected as the points for starting the DIC analysis. The full-field matching results detected by the DIC analysis, in terms of motion of the pixels between two images, are illustrated in Figure 5(E and F). It should be pointed out that many points in the reference image do not have their matching points in the target image because they have moved out of the image boundaries.

Figure 6 demonstrates another simulation, where the reference and target images have periodic patterns. Because there are many nearly identical or very similar regions in the images, the existing relevant techniques typically fail in performing automated initial guess for the DIC analysis. With the presented SIFT and iRANSAC algorithms, the initial guess issue can be easily coped with. A representative result is shown in Figure 6(C and D), where 44 matching pairs of points has been detected and three of them have been selected for the initial guess task. The asymmetry of those point locations is related to the noise and the random function used in the algorithm. The full-field matching results detected by the subsequent DIC analysis are illustrated in Figure 6(E and F).

A simple experiment involving large translation and rotation was conducted to test the capability of the

proposed approach. The two captured images are shown in Figure 7(A and B), and the 134 matching pairs detected by the proposed technique are highlighted in Figure 7(C and D). A careful examination on the matching pairs verified the correctness of the correspondences. The results of the full-field DIC analysis following the automated initial guess are presented in Figure 7(E and F).

The proposed technique was also applied to the DIC analysis of images captured during a three-point bending test of a specimen made of flexible material. Figure 8 shows two representative images and their relevant results. It is noteworthy that because this particular specimen is nearly symmetric, the best matching points can be found in either the left or the right region, but not both. The actual matching detection result depends on the performance of the random function used in the iRANSAC processing. Furthermore, in this particular application, either of the ends will dominate in the SIFT-iRANSAC analysis with regard to the number of final matching points, and the centre region will not be selected by the iRANSAC process because it has a less number of matching points there. In other words, the SIFT process can detect numerous matching pairs everywhere, including the centre region. However, the iRANSAC process will select the group containing the largest number of matching pairs that satisfy Equation (3). Figure 8(C and D) show the matching points in both regions obtained by two separated operations, and the one that yielded matching in the right region was adopted without preference for the subsequent DIC analysis. The final DIC analysis results shown in Figure 8(E and F) help demonstrate the applicability of the proposed scheme for automated fast DIC initial guess.

It is noteworthy that it took less than 1 s to accomplish the automated initial guess for each of the aforementioned simulation and real tests, which involve images varying from 512×512 pixels to 1600×900 pixels. Such processing time is quite short with respect to the entire DIC analysis time. In addition, the final DIC analysis in the two simulations yielded high-accuracy results. The accuracy of the DIC analysis is not presented because it is affected by many factors and is beyond the scope of this paper.

Discussion

There are a few parameters that should be set for the proposed SIFT-iRANSAC technique for the DIC initial guess calculation. They include the initial value of the Gaussian standard deviation (e.g. 1.6), the number of octaves (e.g. 6), the number of layers in each octave (e.g. 6), the FLANN distance ratio (e.g. 0.8), the minimum distance between accepted matching points (e.g. 1 pixel), the requirement on the minimum number of detected matching pairs (e.g. 30), etc. In practice, there is generally no need to specify these parameters because the default parameters, supported by numerous testing under various circumstances, always work well for the DIC applications. Because the parameters



Figure 8: Analysis of real experimental images. (A) is the reference image and (B) is the target image. (C) and (D) are the detected matching pairs of points in the reference image and the target image, respectively; the adopted three matching pairs are surrounded by dotted circles. (E) and (F) are the digital image correlation-detected motion of pixels in the horizontal and vertical directions, respectively

for the SIFT-iRANSAC analysis are set together with the DIC analysis parameters prior to the entire processing and all the parameters often adopt the default values, the entire DIC analysis is an automated process.

The numbers of the detected matching points and the intermediate matching points obtained during the SIFTiRANSAC process depend on a couple of factors, such as the image size, the speckle pattern details and the processing parameters (the default values mentioned previously are often adopted). To provide a brief overview, for instance, the process yields 1446 and 915 initial feature points for the two images (800 × 800 pixels) shown in Figure 7(A and B), respectively; the numbers of the FLANN matching pairs is 214, and the number of the final iRANSAC-selected matching pairs is 134. The relatively small number of the final matching pairs originates from the objective to obtain the most reliable matching.

There are traditionally a few measures that can help perform automated initial guess for the large rotation and deformation cases in the DIC analysis [14, 15], which include but not limited to capturing many intermediate images during measurement, making easy detectable special markers, applying prior knowledge of the experiment to the analysis and so on. Unlike these schemes, the proposed approach provides an alternative and generic way to cope with the automated DIC initial guess issue. Since it is a once and for all solution, it should be very practical for broad scientific and engineering applications.

Finally, it should be noted that the proposed approach cannot cope well with the periodic patterns that are perfectly identical everywhere. This is fortunately not a problem in real-world applications, where imperfection can provide the small yet desired pattern variations.

Conclusion

In conclusion, by integrating the SIFT and iRANSAC algorithms into the DIC technique, the proposed approach has the ability to accurately and automatically detect a number of matching points from two images even though the involved deformation and rotation are large and/or the images have periodic and identical patterns. These matching points can then be used to provide fast and fully automated initial guess to the DIC analysis. Such a function is highly demanded by the DIC technique. The validity of the proposed approach has been demonstrated by both computer simulation and real experiment.

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation (NSF) under grant 0825806 and the United States Army Research Office (USARO) under grant W911NF-10-1-0502.

References

- 1. Sharpe, W. (2008) Handbook of Experimental Solid Mechanics. Springer, New York.
- 2. Sutton, M., Orteu, J. and Schreier H. (2009) Image Correlation for Shape, Motion and Deformation Measurements. Springer, New York.
- Pan, B., Xie, H. and Wang, Z. (2010) Equivalence of digital image correlation criteria for pattern matching. *Appl. Optics* 49, 5501–5509.
- 4. Luu, L., Wang, Z., Vo, M., Hoang, T. and Ma, J. (2011) Accuracy enhancement of digital image correlation with B-spline interpolation. *Opt. Lett.* **36**, 3070–3072.
- 5. Wang, M., Wang, H. and Cen, Y. (2009) High-speed digital image correlation method. *Opt. Lett.* **34**, 1955–1957.
- Wang, Z., Hoang, T., Nguyen, D., Urcinas, A. and Magro, J. (2010) High-speed digital image correlation method: comment. *Opt. Lett.* 35, 2891.
- Barazzetti, L., Scaioni, M. (2010) Development and implementation of image-based algorithms for measurement of deformations in material testing. *Sensors* 10, 7469–7495.
- 8. Zhou, Y., Pan, B. and Chen, Y. (2012) Large deformation measurement using digital image correlation a fully automated approach. *Appl. Optics* **51**, 7674–7683.
- 9. Lowe, D. (2004) Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **60**, 91–110.
- Muja, M. and Lowe, D. (2009) Fast approximate nearest neighbors with automatic algorithm configuration, International Conference on Computer Vision Theory and Applications (VISAPP'09) 24, 331–340.
- 11. Fischler M. and Bolles R. (1981) Random sample consensus a paradigm for model fitting with applications to image analysis and automated cartography. *Comm ACM* **24**, 381–395.
- 12. Chum, O. and Matas, J. (2005) Matching with PROSAC progressive sample consensus, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 220–226.
- 13. Ni, K., Jin, H. and Dellaert, F. (2009) GroupSAC: efficient consensus in the presence of groupings, IEEE 12th International Conference on Computer Vision, 2193–2200.
- 14. Le Cam, J. (2012) A review of the challenges and limitations of full-field measurements applied to large heterogeneous deformations of rubbers. *Strain* **48**, 174–188.
- Barranger, Y., Doumalin, P., Dupre, J. and Germaneau, A. (2012) Strain measurement by digital image correlation: influence of two types of speckle patterns made from rigid or deformable marks. *Strain* 48, 357–365.