

Mesure du champ de déplacement 3D par corrélation d'images numériques photométrique (PhDIC) approche matrix-free avec différenciation automatique

L. Person¹, J.E. Dufour¹, J.C. Passieux¹, J.N. Périé¹

¹ Institut Clément Ader, Université de Toulouse, CNRS/INSA/UT/ISAE-SUPAERO/IMT Mines-Albi, Toulouse, 31400, France, {lperson, passieux, jedufour}@insa-toulouse.fr, jean-noel.perie@iut-tlse3.fr

Résumé —

La corrélation d'images permet la mesure d'un champ de déplacement et repose sur la mise en correspondance de deux images prises à deux instants différents ce qui implique certaines limites. Nous proposons une méthode de mesure de déplacement basée sur les images. Elle repose sur l'inversion du problème direct de génération d'images à partir du chemin de la lumière dans une scène. Une approche sans l'assemblage de matrices a été développée afin d'obtenir des temps de calcul acceptables. La méthode est illustrée sur un essai de torsion, pour lequel un déplacement 3D est mesuré à partir d'une seule vue.

Mots clés — Corrélation d'images Photométrique, Différenciation automatique, Gradient conjugué Matrix-free.

1 Contexte

Le développement récent des technologies d'imagerie a permis l'essor des méthodes de mesure de champ complet. La corrélation d'images numérique (DIC) est une méthode de mesure de champs de déplacement [1, 2]. Il s'agit d'un problème inverse basé sur la comparaison de l'image d'une éprouvette dans son état de référence avec une image de cette même éprouvette dans son état déformé. La DIC peut être effectuée à l'aide d'une seule caméra, généralement pour obtenir un champ de déplacement bidimensionnel sur une surface plane, parallèle au plan de la caméra. La mesure DIC peut aussi être réalisée à l'aide de plusieurs caméras (S-DIC) afin d'obtenir une mesure en trois dimensions. D'après le principe de la stéréovision, la condition nécessaire pour obtenir une mesure en 3D est de disposer d'au moins deux points de vue.

La méthode utilisée pour traiter le problème de la S-DIC est une adaptation de la DIC [3, 4], mais elle présente certaines limites. L'une des principales concerne les grands déplacements ou l'occultation : si un point de la surface disparaît du champ de vision d'une caméra, les informations de référence deviennent inutilisables, rendant la correspondance image de référence/image déformée impossible. Une autre limitation concerne les grandes déformations. Les modèles classiques de corrélation d'images ignorent le processus de formation des pixels et utilisent des interpolations pour estimer les niveaux de gris. En cas de déformations importantes, l'image de référence diffère trop de l'image déformée, ce qui peut provoquer des erreurs significatives ou l'échec de l'algorithme.

2 Méthode

Afin de dépasser les limites associées à la S-DIC, nous proposons d'adopter un formalisme différent, dans lequel nous cherchons à comparer une image réelle I_t , acquise à l'instant t , avec une image virtuelle \hat{I}_t , obtenue par rendu graphique et correspondant au même instant t . Cette image virtuelle est générée à l'aide d'un algorithme de rendu physiquement réaliste, permettant de simuler la propagation de la lumière au travers de la scène.

2.1 Rendu basé sur la physique

Le rendu physique vise à générer des images en simulant le comportement réel de la lumière dans une scène, en se basant sur l'optique géométrique [5, 6]. Une image se forme lorsqu'un capteur enregistre la lumière après qu'elle a été émise par des sources lumineuses et réfléchiée par les objets.

Intuitivement, une image est produite lorsque la lumière émise par une source, puis réfléchiée par les surfaces des objets observés, est captée par un capteur. Ainsi, dans les algorithmes de rendu basés sur la physique [5], avant même de définir les équations, la scène est décrite. Une scène est composée des éléments suivants :

- *Sources lumineuses* qui émettent une radiance dans la scène.
- *Formes géométriques* représentant les objets 3D de la scène. Ces formes sont modélisées sous forme de maillages surfaciques 3D T3.
- *Caractéristiques de surface* : chaque forme possède une couleur et des propriétés de réflexion spécifiques.
- *Capteur* qui reçoit et mesure la radiance incidente de la scène. Dans cette étude, le capteur est une caméra, mais il peut représenter tout appareil de mesure de la lumière.

En pratique, les algorithmes de rendu basés sur la physique produisent des mesures de grandeurs radiométriques (comme la radiance). L'équation de mesure décrit le rôle du capteur dans la formation de ces mesures :

$$I = \int_{\mathcal{A}} \int_{S^2} W_e(\underline{x}, \underline{\omega}) L_i(\underline{x}, \underline{\omega}) dA(\underline{x}) d\sigma^\perp(\underline{\omega}), \quad (1)$$

où \mathcal{A} désigne la surface du capteur, S^2 la surface d'une sphère unité, et W_e la fonction de réponse du capteur en position \underline{x} et dans la direction $\underline{\omega}$. L_i représente la radiance incidente provenant de la scène. Enfin, $d\sigma^\perp$ correspond à l'angle solide mesuré en \underline{x} , projeté sur la sphère unité.

La principale difficulté de cette équation réside dans l'obtention de la radiance incidente L_i . Pour déterminer la radiance reçue par le capteur, il est nécessaire de simuler le transport de la lumière dans la scène. Nous faisons l'hypothèse que le milieu de propagation de la lumière est négligeable. Ainsi, la lumière peut être considérée comme se déplaçant en ligne droite, ce qui permet de définir l'équation de transport :

$$L_i(\underline{x}, \underline{\omega}) = L_o(\underline{r}(\underline{x}, \underline{\omega}), -\underline{\omega}), \quad (2)$$

où \underline{r} désigne la fonction de lancer de rayons, qui retourne le point d'intersection le plus proche entre un rayon d'origine \underline{x} et de direction $\underline{\omega}$. La radiance sortante L_o peut alors être calculée comme suit :

$$L_o(\underline{x}, \underline{\omega}_o) = L_e(\underline{x}, \underline{\omega}_o) + \int_{S^2} f(\underline{x}, \underline{\omega}_i, \underline{\omega}_o) L_o(\underline{r}(\underline{x}, \underline{\omega}_i), -\underline{\omega}_i) d\sigma^\perp(\underline{\omega}_i), \quad (3)$$

où L_e représente la radiance émise. L'intégrale calcule la radiance incidente en un point \underline{x} , provenant de toutes les directions $\underline{\omega}_i$, puis réfléchiée dans la direction $\underline{\omega}_o$. La fonction f correspond à la fonction de distribution bidirectionnelle de réflexion (BRDF). Dans cette étude, nous adoptons un modèle simplifié de BRDF, en supposant que toutes les surfaces observées sont Lambertiennes :

$$f(\underline{x}, \underline{\omega}_i, \underline{\omega}_o) = \frac{\rho(\underline{x})}{\pi}, \quad (4)$$

où ρ correspond à l'albedo de la surface touchée par un rayon lumineux.

2.2 Rendu différentiable

Au final, notre objectif est de formuler un problème inverse et de calculer les gradients du rendu \hat{I} obtenu à l'aide d'un algorithme de rendu basé sur la physique. Étant donné la complexité des équations de rendu, la *différenciation automatique* (DA) constitue une solution particulièrement adaptée.

Contrairement à d'autres techniques, la différenciation automatique ne repose pas sur des approximations numériques pour estimer les dérivées. Elle s'appuie sur la règle de la chaîne et permet de calculer la dérivée d'une fonction $y \circ w_{n-1} \circ w_{n-2} \circ w_{n-3}$ par rapport à une variable x selon deux modes distincts :

$$\left\{ \begin{array}{l} \text{Accumulation directe : } \frac{\partial y}{\partial x} = \frac{\partial y}{\partial w_{n-1}} \left(\frac{\partial w_{n-1}}{\partial w_{n-2}} \left(\frac{\partial w_{n-2}}{\partial w_{n-3}} \frac{\partial w_{n-3}}{\partial x} \right) \right) \\ \text{Accumulation inverse : } \frac{\partial y}{\partial x} = \left(\left(\frac{\partial y}{\partial w_{n-1}} \frac{\partial w_{n-1}}{\partial w_{n-2}} \right) \frac{\partial w_{n-2}}{\partial w_{n-3}} \right) \frac{\partial w_{n-3}}{\partial x} \end{array} \right. . \quad (5)$$

Ainsi, grâce à une implémentation d'un algorithme de rendu basé sur la physique combiné à la différenciation automatique, il est possible de calculer la dérivée du rendu \hat{I}_t par rapport à n'importe quelle variable d'intérêt. En particulier, nous pouvons calculer :

$$\left\{ \begin{array}{l} \text{Accumulation directe : } \underline{u} = \underline{J}_{\hat{I}_t, \lambda} \underline{v} \\ \text{Accumulation inverse : } \underline{v} = \underline{J}_{\hat{I}_t, \lambda}^T \underline{u} \end{array} \right. , \quad (6)$$

où $\underline{J}_{\hat{I}_t, \lambda}$ désigne la matrice jacobienne du rendu \hat{I}_t par rapport à la variable λ . En définitive, la différenciation automatique permet de minimiser des fonctions objectif de la forme suivante :

$$\mathcal{L}(\hat{I}_t(\lambda)) \in \mathbb{R}. \quad (7)$$

2.3 Fonctionnelle photométrique

Après avoir présenté la résolution du problème direct de génération d'images et la manière dont nous pouvons calculer ses dérivées par rapport à n'importe quelle variable, il est nécessaire de définir la fonctionnelle associée au problème inverse que nous souhaitons résoudre.

Pour rappel, notre objectif est de mesurer un champ de déplacement \underline{u}_t . Cependant, si l'on analyse le problème, en supposant que la source lumineuse, la forme et la position de l'objet observé ainsi que la caméra sont connues, une inconnue subsiste : les caractéristiques de la surface observée. Sous l'hypothèse Lambertienne, cette inconnue correspond à l'albédo ρ de la surface. Comme nous souhaitons également l'estimer, nous formulons une fonctionnelle qui dépend à la fois du déplacement \underline{u}_t et de l'albédo ρ .

Ainsi, la *fonctionnelle photométrique* s'écrit [7] :

$$\mathcal{F}((\underline{u}_t)_t, \rho) = \sum_{t=1}^{N_{frame}} \|\underline{I}_t - \hat{I}_t(\underline{u}_t, \rho)\|_p^p, \quad (8)$$

où N_{frame} représente le nombre de pas de temps, \underline{I}_t l'image réelle au pas de temps t , et p la norme utilisée sur le domaine de l'image (en pratique, $p = 1$ ou $p = 2$ selon le problème). Cette fonctionnelle exploite directement l'intensité lumineuse mesurée par chaque pixel. Les inconnues à estimer sont donc une collection de champs de déplacement et l'albédo.

La minimisation de cette fonctionnelle est réalisée par minimisation alternée, résolvant successivement deux sous-problèmes :

- *Estimation des déplacements*, en supposant que l'albédo ρ est connu [7]. Cette étape correspond à la corrélation d'images numériques photométrique (PhDIC). Il est important de noter que la mesure de déplacement peut être faite pour n'importe quel point visible dans l'image i , peu importe si ce point est visible dans la configuration de référence.
- *Estimation de l'albédo*, en supposant que la collection de champs de déplacement $(\underline{u}_t)_t$ est connue. On parle de Super-Résolution (SR) [8].

Une fois l'albédo estimé, il peut être réutilisé pour mesurer n'importe quel champ de déplacement, sous n'importe quelle condition d'illumination ou de chargement. En pratique il est donc possible de mesurer un déplacement 3D à partir d'une seule vue, néanmoins la fonctionnelle comporte beaucoup de minima locaux. Le choix d'une réduction de la base cinématique adaptée est donc primordial pour une convergence de l'algorithme.

2.4 Implémentation

Dans cette section, nous nous concentrons principalement sur l'algorithme d'estimation du champ de déplacement \underline{u}_t , en supposant que l'albédo ρ est connu, car c'est essentiellement la mesure du champ

de déplacement qui nous intéresse. Pour ce faire, nous minimisons la fonctionnelle \mathcal{F} avec $p = 2$, ce qui correspond à un problème de minimisation d'une norme des moindres carrés non linéaire.

La recherche de \underline{u}_t est effectuée de manière itérative à l'aide de l'algorithme de Gauss-Newton, selon la relation : $\underline{u}_t^{(i+1)} = \underline{u}_t^{(i)} + d\underline{u}_t^{(i)}$. En développant \mathcal{F} autour de $\underline{u}_t^{(i)}$, chaque itération revient à résoudre le problème linéaire suivant :

$$\text{Trouver } d\underline{u}_t^{(i)} \in (L^2(\Omega))^3, \quad \underline{\underline{J}}_{\hat{I}_t^{(i)}}^T \underline{\underline{J}}_{\hat{I}_t^{(i)}} d\underline{u}_t^{(i)} = \underline{\underline{J}}_{\hat{I}_t^{(i)}}^T \left(I_t - \hat{I}_t^{(i)} \right), \quad (9)$$

où $\underline{\underline{J}}_{\hat{I}_t^{(i)}}$ représente la jacobienne du rendu par rapport aux positions des nœuds du maillage surfacique T3, et Ω désigne le domaine de définition du maillage. À chaque itération, il est donc nécessaire de générer un rendu $\hat{I}_t^{(i)}$ et de calculer la jacobienne correspondante $\underline{\underline{J}}_{\hat{I}_t^{(i)}}$. Naïvement si on utilise les différences finies pour calculer cette matrice, il faudrait générer $N_{ddl} + 1$ rendus. Même si la génération du rendu direct est relativement rapide, le calcul explicite de la matrice jacobienne est extrêmement coûteux, voire rédhibitoire, en particulier quand la cinématique est paramétrée par quelques milliers de degrés de liberté éléments finis. Il est donc nécessaire d'adopter une approche permettant la résolution du système linéaire de chaque itération dans un temps de calcul acceptable, c'est-à-dire de l'ordre de la seconde.

Pour pallier cette difficulté, l'idée consiste à développer une approche sans matrice, basée sur un solveur itératif de type Krylov, en l'occurrence l'algorithme du gradient conjugué. Ces systèmes linéaires étant imbriqués dans un solveur moindres carrés non-linéaire global, nous montrons qu'il n'est pas utile de converger cet algorithme : le nombre maximal d'itérations est fixé à n_{ite} , en pratique inférieur à 10. Cette approche peut légèrement augmenter le nombre d'itérations de Gauss-Newton nécessaires pour atteindre la convergence, mais elle réduit drastiquement le temps de calcul global.

Un autre avantage majeur de cette méthode est qu'elle nécessite uniquement des produits matrice-vecteur. Ainsi, grâce aux propriétés de la différenciation automatique, nous pouvons résoudre le problème sans jamais former explicitement la matrice jacobienne $\underline{\underline{J}}_{\hat{I}_t^{(i)}}$.

L'algorithme de gradient conjugué modifié utilisé est présenté ci-après.

Gradient conjugué Matrix-free

- 1: **Données** : Image de référence I_t , image générée par rendu $\hat{I}_t^{(i)}$
- 2: **Initialisation** : Choisir $d\underline{u}_t^{(i),0}$ (= 0 en pratique)
- 3: Calcul du RHS par DA et en mode *accumulation directe* : $\underline{b} = \underline{\underline{J}}_{\hat{I}_t^{(i)}}^T \left(I_t - \hat{I}_t^{(i)} \right)$
- 4: Calcul du LHS par DA en mode *accumulation directe* puis *inverse* :

$$\underline{\underline{H}} d\underline{u}_t^{(i),0} = \underline{\underline{J}}_{\hat{I}_t^{(i)}}^T \left(\underline{\underline{J}}_{\hat{I}_t^{(i)}} d\underline{u}_t^{(i),0} \right)$$
- 5: Calculer le résidu initial : $\underline{r}^0 = \underline{b} - \underline{\underline{H}} d\underline{u}_t^{(i),0}$
- 6: Calcul de la direction de recherche initiale $\underline{d}^0 = \underline{r}^0$
- 7: **Pour** $k = 1$ à n_{ite} **faire**
- 8: Calcul de $\underline{\underline{H}} \underline{d}^k$ par DA et *accumulation directe* puis *inverse* :

$$\underline{\underline{H}} \underline{d}^k = \underline{\underline{J}}_{\hat{I}_t^{(i)}}^T \left(\underline{\underline{J}}_{\hat{I}_t^{(i)}} \underline{d}^k \right)$$
- 9: Calcul du pas optimal : $\alpha^k = \frac{\underline{r}^{kT} \underline{r}^k}{\underline{d}^{kT} \underline{\underline{H}} \underline{d}^k}$
- 10: Calcul de l'itéré : $d\underline{u}_t^{(i),k+1} = d\underline{u}_t^{(i),k} + \alpha^k \underline{d}^k$
- 11: Calcul du résidu : $\underline{r}^{k+1} = \underline{r}^k - \alpha^k \underline{\underline{H}} \underline{d}^k$
- 12: Calcul du paramètre d'orthogonalisation : $\beta^{k+1} = \frac{\underline{r}^{k+1T} \underline{r}^{k+1}}{\underline{r}^{kT} \underline{r}^k}$
- 13: Mise à jour de la direction de recherche : $\underline{d}^{k+1} = \underline{r}^{k+1} + \beta^{k+1} \underline{d}^{k+1}$
- 14: **Fin pour**
- 15: **Sortie** : Incrément de déplacement $d\underline{u}_t^{(i),n_{ite}}$

3 Torsion de grande amplitude

La méthode présentée précédemment permet donc de mesurer un déplacement en trois dimensions à partir d'un unique point de vue. Nous illustrerons cette approche sur un cas de grande déformation, ici une torsion de forte amplitude. Nous étudierons donc la torsion d'un tube d'adhésif thermoplastique extrudé. Ce tube présente un diamètre de 11.5 mm et une longueur observable de 142.4 mm entre les mors de la machine de torsion. Le mouchetis a été appliqué à l'aide d'une bombe de peinture. Les images ont été acquises avec un capteur monochromatique et présentent une résolution de 2048×2048 pixels. Deux bases de données d'images ont été capturées :

1. *Base de données de scan* : la méthode de mesure du déplacement repose sur la connaissance de l'albédo, rendant l'étape d'estimation de celui-ci cruciale. Pour simplifier cette estimation, nous avons choisi d'acquérir des images correspondant à des mouvements de corps rigide, ici des rotations. Cela améliore l'efficacité de l'algorithme en permettant de réduire l'espace cinématique. La base de données de scan est composée de $N_{frame} = 62$ images, avec un pas de rotation de 5.65° autour de l'axe vertical.
2. *Base de données de mesure* : cette base constitue l'ensemble d'images utilisé pour mesurer la cinématique d'intérêt. Elle comprend $N_{frame} = 72$ images correspondant à la torsion du tube de colle. La base inférieure du tube est fixée, et une rotation jusqu'à plus de 360° est imposée autour de l'axe vertical sur le mors supérieur.

La figure 1 montre quelques images de l'essai de torsion, avec en rouge l'évolution d'une ligne de l'éprouvette au fil des images. Cet essai complexe met en évidence plusieurs défis liés à la mesure du champ de déplacement. Tout d'abord, une seule vue est disponible à chaque instant pour mesurer un déplacement 3D. De plus, la torsion est de grande amplitude, entraînant de fortes déformations dans certaines zones de l'éprouvette et surtout des problèmes d'auto-occultation qui rendent la DIC classique quasiment inopérante. Enfin, la zone d'intérêt est peu résolue, ce qui complique l'extraction précise des déplacements.

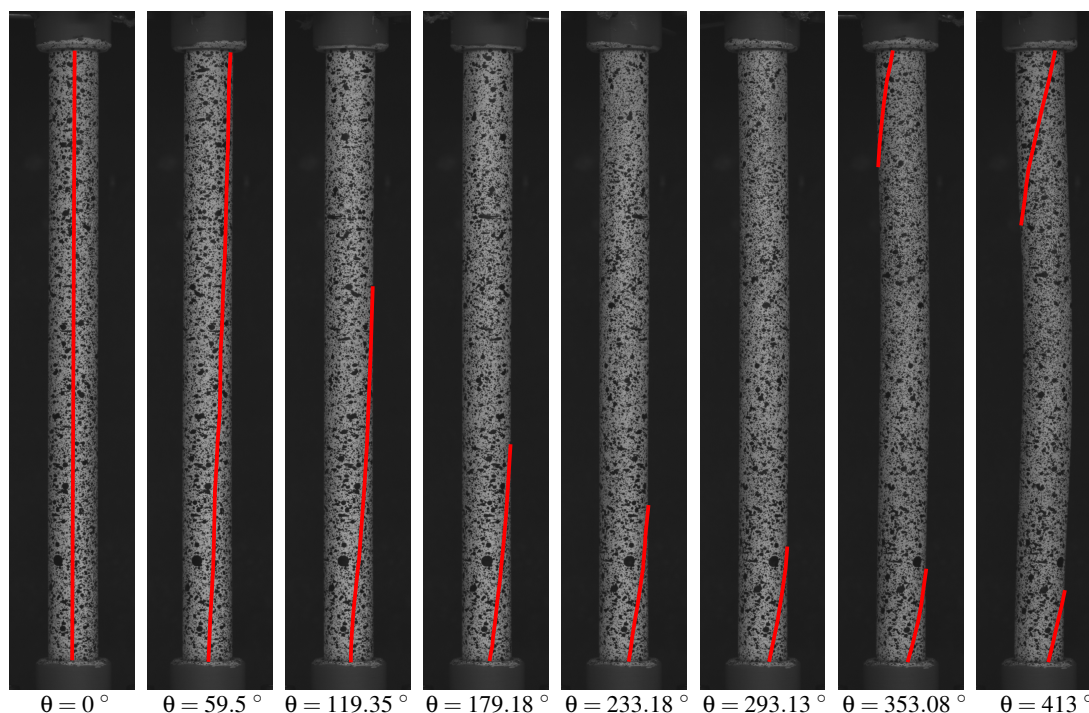


FIGURE 1 – Extrait de la base de données d'images de torsion avec représenté en rouge l'évolution au cours des images de la visibilité de cette ligne

Nous montrerons que par notre algorithme il est possible de mesurer un albedo de toute la surface visible de l'éprouvette, ce qui nous permettra de construire un jumeau numérique. Nous verrons qu'il est possible de mesurer la torsion de cette éprouvette à partir d'un seul point de vue. Enfin, nous illustrerons

les différents coûts de calcul associés à cette méthode et démontrerons l'intérêt du gradient conjugué matrix-free afin d'obtenir une mesure dans un temps acceptable.

4 Conclusion

En combinant des outils issus de la vision par ordinateur, du rendu basé sur la physique nous avons pu formuler le problème direct de génération d'images lié au problème inverse de mesure des champs de déplacement à partir d'images. Cette approche permet une mesure directe de tout point visible, sans besoin de le voir dans la configuration de référence, tout en prenant en compte les propriétés d'éclairage de la scène et en dépassant les limitations de la DIC/S-DIC face aux grandes transformations et l'occlusion. L'algorithme PhDIC repose sur une méthode itérative de type Gauss-Newton. À chaque itération, la différenciation automatique permet de résoudre efficacement les systèmes linéaires via un gradient conjugué non convergé, sans avoir besoin d'assembler les matrices du système, dans des temps humains raisonnables.

Références

- [1] T. C. Chu, W. F. Ranson, M. A. Sutton. Applications of digital-image-correlation techniques to experimental mechanics, *Experimental Mechanics*, Volume 25 (3) (1985) 232-244.
- [2] F. Hild, S. Roux. Digital Image Correlation : from Displacement Measurement to Identification of Elastic Properties - a Review, *Strain*, Volume 42 (2) (2006) 69-80.
- [3] B.D. Lucas, T. Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, Volume 2 (1989).
- [4] J.E. Pierré, J.C. Passieux, J.N. Périé. Finite element stereo digital image correlation : framework and mechanical regularization. *Experimental Mechanics*, Volume 57 (2017) 443-456.
- [5] M. Pharr, W. Jakob, G. Humphreys. Physically based rendering : From theory to implementation, *MIT Press*, 2023.
- [6] R. Szeliski. Computer vision : algorithms and applications, *Springer Nature*, 2022.
- [7] R. Fouque, R. Bouclier, J.C. Passieux, J.N. Périé. Photometric DIC : a unified framework for global Stereo Digital Image Correlation based on the construction of textured digital twins. *Journal of Theoretical, Computational and Applied Mechanics*, (2022) 1-32.
- [8] S. Farsiu, M.D. Robinson, M. Elad, P. Milanfar. Fast and robust multiframe super resolution. *IEEE transactions on image processing*, Volume 13 (10) (2004) 1327-1344.