

Projection de texture sur des objets 3D réels

Thierry MOLINIER¹, David FOFI¹, Patrick GORRIA¹, Joaquim SALVI², Fabrice MERIAUDEAU¹

¹Laboratoire d'Electronque, Informatique et Image
12 rue de la Fonderie 71200 Le Creusot, France

²Laboratoire Vicorob
Université de Gérone, Gerone, Espagne

tmolinier@iutlecreusot.u-bourgogne.fr, dfofi@iutlecreusot.u-bourgogne.fr,
pgorria@iutlecreusot.u-bourgogne.fr, qsalvi@silver.udg.es, fmeriaudeau@iutlecreusot.u-bourgogne.fr

Résumé – Nous proposons une méthode pour projeter une texture sur un objet réel sans aucun contact. Une telle technique peut être utile dans les domaines médicales ou archéologiques. La technique de lumière structurée est une technique optique qui permet l'acquisition des formes. En projetant une mire adéquate sur la surface d'un objet et en capturant la scène, de nombreux points de correspondance peuvent être estimés et la reconstruction 3D par triangulation est alors possible. Une fois l'objet reconstruit et donc sa position connue, la projection de texture sur cet objet est alors possible.

Abstract – We propose a technique to project texture on real object without any contact with it. Such techniques can be apply on medical application or archeologic. Coded structured light is an optical technique based on active stereovision which allows shape acquisition. By projecting a suitable set of light pattens onto the surface of an object and capturing images with a camera, a large number of correspondences can be found and 3D points can be reconstructed by means triangulation. Once the 3D reconstruction is done and then the pose of the real object is estimated, the projection of texture on the surface of this object is possible.

1 Introduction

La reconstruction de surfaces est l'un des thèmes les plus importants dans la vision artificielle de part son large champ d'application. Quelques exemples en sont la rétro-conception, l'inspection industrielle de produits manufacturés, la reconnaissance de formes. Parmi toutes les techniques de reconstruction [6], la stéréovision est basée sur l'observation d'une scène à partir d'au moins deux points de vues et l'acquisition de points de correspondance entre ces vues. Une fois la correspondance établie, la triangulation de ces points, ie l'estimation de leur position dans un repère 3D, est possible, si les caméras ont été préalablement étalonnées [4],[8]. Trouver des points de correspondance se révèle une tâche ardue, même en prenant en compte des contraintes épipolaires. Pour contourner cette difficulté, la technique de la lumière structurée consiste à remplacer une des caméras par un vidéoprojecteur qui projette une mire sur la scène ou l'objet qui nous intéresse.

La projection de texture convertit une texture bidimensionnelle en une texture tridimensionnelle en la projetant dans une ou plusieurs directions. D'habitude, la texture est projetée sur un objet virtuel dans un monde virtuel avec de nombreuses applications, comme la réalité augmentée. En revanche, la projection de texture dans le monde réel n'est que peu explorée dans la communauté du traitement d'image, malgré ses nombreuses applications potentielles. Habituellement on projette une texture sur un objet réel pour faciliter le problème de mise en correspondance pour l'acquisition de surface, par exemple sur des objets dépourvus de texture, par la projection de mire codée ou

pseudo-codée. Néanmoins peu d'articles traitent principalement de la projection de texture pour appliquer de manière adéquate une texture tridimensionnelle sur un objet tridimensionnel réel. Les applications d'une telle technique sont variées et intéressantes, qui permettent en particulier de percevoir l'apparence visuelle de l'objet avant même de lui appliquer la texture de manière plus conventionnelle. On peut citer en particulier les musées virtuels, le design vestimentaire ou le cinéma.

En connaissant la surface et par la suite les normales de ces surfaces, la projection de texture peut servir à accroître la segmentation des couleurs sur un objet réel ou bien encore à appliquer une texture sur un objet justement dépourvu de toute texture. L'emploi que nous faisons du vidéoprojecteur est donc double :

- une utilisation active est faite lors de la projection d'une mire afin de faciliter la mise en correspondance des vues.
- une utilisation passive est en revanche faite lors de la projection de la texture sur l'objet.

Notre article se découpe en trois grandes parties : la première partie traite de l'étalonnage du projecteur, la seconde décrit la technique pour trouver les points de correspondance et la dernière partie montre les résultats obtenus. La conclusion permettra de voir les perspectives de nos travaux.

2 Etalonnage du système

Nous avons créé un capteur composé de deux caméras et d'un vidéoprojecteur. Ce système nous offre une grande flexibilité et permet non seulement l'acquisition de formes 3D mais aussi la projection de texture. En pratique il est possible de se limiter à l'emploi des deux caméras pour trouver les points de correspondance et trianguler la scène ou bien d'utiliser une caméra et le projecteur comme système de lumière structurée (dans le cas d'objet sans texture). La géométrie épipolaire doit être estimée. Nous avons privilégié l'emploi des matrices fondamentales plutôt que celui du tenseur trifocal. Bien que le tenseur trifocal [7] soit dédié à l'estimation de la géométrie épipolaire pour trois vues, sa robustesse au bruit est plus faible que celle des matrices fondamentales, qui traduisent la géométrie épipolaire pour deux vues. Nous estimons donc les trois matrices pour chaque binôme de vues parmi les trois.

L'étalonnage fort du système est nécessaire pour pouvoir projeter avec précision une texture virtuelle sur la surface de l'objet. En fait les caractéristiques optiques du projecteur doivent être connues pour pouvoir projeter la texture à l'échelle mais aussi la position relative de l'objet par rapport au projecteur, c'est à dire connaître les paramètres intrinsèques et extrinsèques du projecteur. L'étalonnage nous permet d'estimer les 2 matrices des caméras, respectivement $\mathbf{P}_{c1} \approx \mathbf{K}_{c1} \cdot [\mathbf{R}_{c1} \ \mathbf{t}_{c1}]$ et $\mathbf{P}_{c2} \approx \mathbf{K}_{c2} \cdot [\mathbf{R}_{c2} \ \mathbf{t}_{c2}]$ et la matrice du projecteur $\mathbf{P}_p \approx \mathbf{K}_p \cdot [\mathbf{R}_p \ \mathbf{t}_p]$. Les matrices \mathbf{R}_i et \mathbf{t}_i représentent la position de la caméra, ou du projecteur, par rapport au repère choisi (avec $i = c_1, c_2$ or p).

L'étalonnage du projecteur étant plus compliqué à réaliser que celui des caméras, une méthode d'étalonnage par homographie peut être avantageusement utilisée [2] [1]. Pour rendre cette étape la plus simple possible, nous avons décidé de séparer l'étalonnage du capteur en deux étapes :

- Les caméras sont étalonnées. Pour augmenter la qualité des résultats, nous estimons la distorsion créée par les caméras et nous réalisons une normalisation des données avant le calcul [1].
- une mire est projetée sur un plan et la scène est capturée par les deux caméras. Les points illuminés sont triangulés pour obtenir une correspondance entre un nuage de points 2D et 3D.
- la méthode de Faugeras-Toscani [3] est utilisée pour obtenir les paramètres intrinsèques et extrinsèques du projecteur.

3 Mise en correspondance

Trouver des points de correspondance sur un objet dépourvu de texture se révèle un exercice difficile. Pour résoudre ce problème nous utilisons deux méthodes : la géométrie épipolaire fournit pour un point donné d'une image la ligne épipolaire correspondante sur une autre image où le point de correspondance se trouve, limitant par la même la zone de recherche du point de correspondance. La mire structurée que nous projetons permet d'identifier le point de correspondance le long de la ligne épipolaire. Pour esti-

mer les matrices fondamentales utiles au calcul des lignes, nous avons déjà de nombreux points de correspondance entre les trois vues, les points utilisés lors de l'étalonnage du système. La distorsion de l'image créée par la lentille des caméras ne peut être ignorée. Ainsi les matrices fondamentales doivent être estimées à partir des points corrigés.

3.1 La géométrie épipolaire

La matrice fondamentale F est estimée à partir des points de correspondance de deux images, m_1 et m_2 par l'équation 1 :

$$m_1^t \cdot F \cdot m_2 = 0 \quad (1)$$

La figure 1 décrit la relation entre les points m_1 , m_2 et le point 3D m .

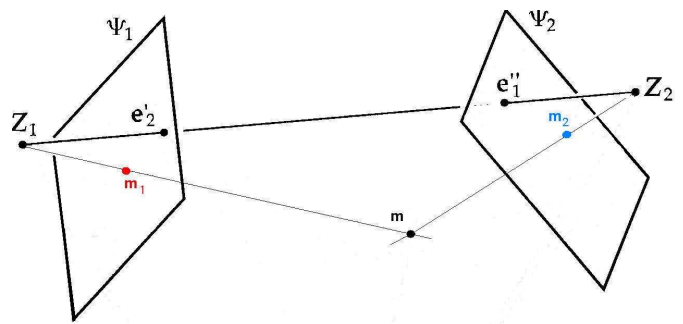


FIG. 1 – Géométrie épipolaire pour deux vues

La matrice fondamentale est estimée pour chaque combinaison de deux vues parmi les trois disponibles. Nous avons choisi de partir de l'image du projecteur, soit la mire structurée, car il est plus facile de reconnaître les éléments de la mire. Armanque [9] a testé près de dix-neuf des méthodes les plus utilisées pour l'estimation des matrices fondamentales. Les résultats expérimentaux démontrent que :

- (a) les méthodes linéaires sont efficaces si les points sont précisément localisés sur l'image et la mise en correspondance aussi.
- (b) les méthodes itératives sont robustes face à un bruit gaussien, mais dégénèrent en présence de points aberrants.
- (c) les méthodes robustes résistent non seulement au bruit mais aussi aux points aberrants.

Ainsi la méthode MLeS obtient les meilleurs résultats alors que le temps d'analyse reste raisonnable. Cette méthode est la plus appropriée pour la détection des points aberrants et ainsi leur élimination.

3.2 Corrélation avec la mire

A partir d'un point de la mire, nous pouvons estimer la ligne épipolaire pour les images des caméras. Le point de correspondance doit se trouver sur cette ligne, ou au moins au voisinage de cette ligne, donc la zone de recherche sera une bande centrée sur cette ligne. La géométrie épipolaire ne peut pas nous donner plus d'information, mais la mire le peut. La mire structurée (voir figure 2) a été créée

de telle sorte que chaque cellule de cinq par cinq carrés élémentaires est unique dans toute l'image.



FIG. 2 – Mire structurée

Trouver les points de correspondance se résume donc maintenant à trouver la même cellule sur deux images. Il existe différentes méthodes pour estimer la corrélation entre une cellule de l'image du projecteur et celle d'une des deux caméras [5], basées sur les filtres d'énergie. Ces filtres sont composés d'un jeu de masques qui caractérisent certaines particularités de texture de l'image. Mais nous allons comparer directement les deux cellules en utilisant la formule 2 :

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\sum_m (A_{mn} - \bar{A})^2} \sqrt{\sum_n (B_{mn} - \bar{B})^2}} \quad (2)$$

\bar{A} et \bar{B} représentant les valeurs moyennes des images A et B .

Nous estimons toute corrélation possible le long de la ligne épipolaire et si le maximum de corrélation est supérieur à un seuil fixé empiriquement (généralement 0.8 sur 1), alors nous considérons avoir trouvé un jeu de points de correspondance.

3.3 Recalage des vues

Nous obtenons deux jeux de points de correspondance, entre le projecteur et les 2 caméras. Malgré le soin apporté à l'étalonnage du système, la triangulation des jeux de points laisse apparaître les erreurs d'étalonnage. Fort heureusement nous obtenons des points de correspondance entre les deux caméras, grâce à notre méthode mais aussi par le fait que certains points de la mire du projecteur servent de points de correspondance pour les deux caméras. Ces points nous permettent de déterminer les corrections à apporter : l'étalonnage du projecteur résultant de celle des caméras, les erreurs de calibration se cumulent donc les points 3D triangulés à partir des caméras sont les moins bruités. Nous nous servons de ces points pour recalibrer les points 3D issus de la triangulation des points de correspondance entre le projecteur et une caméra.

La dernière étape du recalage consiste à recalibrer sur notre reconstruction 3D un modèle 3D complet de l'objet pour pouvoir texturer complètement l'objet du point de vue du projecteur. La figure 3 récapitule toutes les étapes de notre méthode.

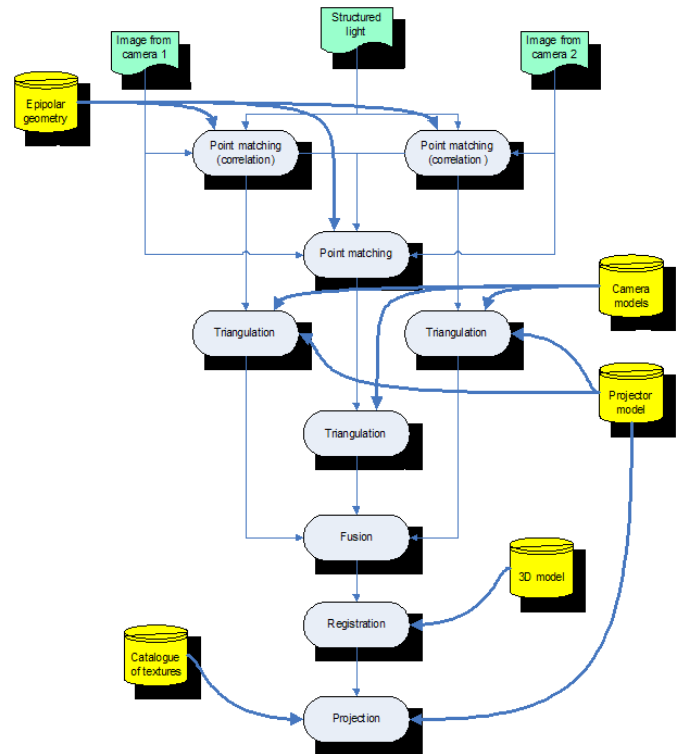


FIG. 3 – diagramme de la méthode

4 Résultats expérimentaux

La figure 4 présente les images des caméras lorsque la mire structurée est projetée, ainsi que les points de correspondance trouvés avec l'image vue par le projecteur, soit la mire structurée. On note que plus l'objet est de biais par rapport au projecteur, plus la mire est déformée sur sa surface et plus la corrélation est faible, donc le nombre de points trouvés diminue.

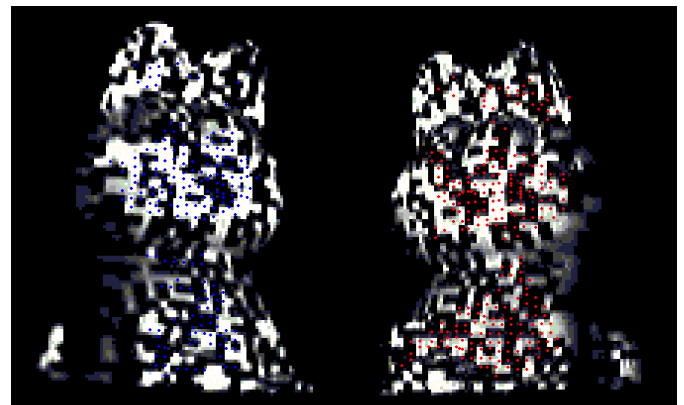


FIG. 4 – Points de correspondance trouvés

En combinant la géométrie épipolaire et la technique de lumière structurée, nous arrivons à obtenir plus de points de correspondance qu'avec un détecteur classique de points d'intérêt pour des objets dépourvus de texture. La figure 5 montre les différents nuages de points recalés dans un repère centré sur le projecteur.

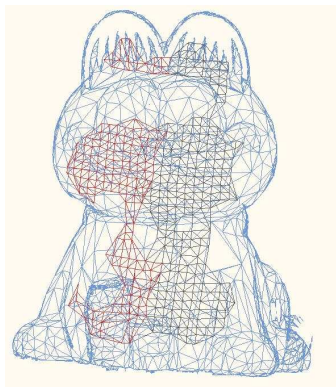


FIG. 5 – Recalage de différents nuages de points

Les figures 6 et 7 exposent enfin une image de l'objet réel, et enfin ce même objet soumis à différentes textures.



FIG. 6 – modèle original



FIG. 7 – modèles texturés

5 Conclusion

La configuration du système, en particulier la position respectives des caméras, ne permet pas d'employer des algorithmes classiques pour déterminer des points de correspondance entre les deux images. Mais l'utilisation active du projecteur, par la projection d'une mire structurée, permet de résoudre cette difficulté. Combiner la mire avec la géométrie épipolaire, en limitant ainsi la zone de recherche dans l'image, permet de réduire le temps de calcul.

Nos futurs travaux devraient se concentrer sur l'acquisition précise et automatique des cellules de la mire sur l'objet et aussi le centre de cette mire, voir l'emploi d'autres types de mires structurées. Cela devrait non seulement accroître le nombre de points de correspondance, en particulier dans les zones à fort relief de l'objet où la mire est fortement déformée, mais aussi sur la précision de la reconstruction 3D, d'où découlera plus d'aisance pour recalculer précisément le modèle de l'objet.

Références

- [1] J.P.Bouguet, <http://www.vision.caltech.edu/bouguetj>, 1995.
- [2] Cotting, D. ; Naef, M. ; Gross, M. ; Fuchs, H., Embedding Imperceptible Patterns into Projected Images for Simultaneous Acquisition and Display., Third IEEE and ACM International Symposium on Mixed and Augmented Reality, 2004.
- [3] Faugeras and Toscani., The Calibration Problem for Stereo, CVPR86, 1986, 15-20.
- [4] O. Faugeras, Three-Dimensional Computer Vision, MIT Press, Cambridge, 1993.
- [5] Garcia Campos, Rafael, A proposal to estimate the motion of an underwater vehicle through visual mosaicking, EIA Electrònica, Informàtica i Automàtica, University of girona, Spain, 2001.
- [6] R. Jarvis., Range sensing for compute vision, Advances in Image Communications, Elsevier Sciences Publishers,1993.
- [7] T. Molinier, D. Fofi, R. Seulin, F. Meriaudeau, Trifocal Tensor Estimation Applied to Imperceptible Structured Light The International Topical Meeting on Optical Sensing and Artificial Vision, Saint Petersburg, Russia, October 2004.
- [8] J. Salvi, X. Armangué, J. Battle., A comparative review of camera calibrating methods with accuracy evaluation, Pattern Recognition, 1617-1635, 35, 2002.
- [9] X. Armangué and J. Salvi, Overall View Regarding Fundamental Matrix Estimation, Image and Vision Computing, volume 21, February 2003, 205-220.