HAL
open science

Rendu à base de vidéos non calibrées

Vincent Nozick, Hideo Saito

To cite this version:

Vincent Nozick, Hideo Saito. Rendu à base de vidéos non calibrées. Journée Scientifique Francophone JSF 2006, Dec 2006, Japon. pp.36. hal-00733836

HAL Id: hal-00733836

https://hal.science/hal-00733836v1

Submitted on 19 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Domaine de recherche: Informatique
Spécialité: Informatique graphique

Rendu à Base de Vidéos non calibrées

Vincent Nozick, Hideo Saito
Hideo Saito Laboratory,
Universite de Keio,
3-14-1 Hiyoshi, Kohoku-ku
223-8522, Japan
E-mail: nozick@ozawa.ics.keio.ac.jp

Résumé:

Étant donné un ensemble de caméras fixes filmant une même scène, le rendu à base de vidéos consiste à générer des nouvelles images de cette scène à partir de nouveaux points de vue. L'utilisateur a ainsi l'impression de pouvoir déplacer une caméra virtuelle dans la scène alors qu'en réalité, toutes les caméras sont fixes. Le rendu à base de vidéos permet aussi de générer une reconstruction 3d de la scène filmée.
Nos recherches portent d'une part sur la possibilité de générer des nouvelles vues de la scène en temps réel et en direct (par opposition au rendu en différé). D'autre part, nous orientons nos recherches vers l'emploi de caméras non calibrées, c'est-à-dire de caméras dont on ne connaît pas les caractéristiques techniques (focale, position, orientation, ...). L'utilisation de caméras non calibrées rend l'utilisation de ce genre de systèmes beaucoup moins contraignante. Par ailleurs, nous essayons de restreindre le matériel utilisé à du matériel “grand public” comme des webcams par exemple afin de rendre ce système accessible à tous.

Abstract:

Given a small set of video streams of a scene, video-based rendering techniques compute new views of this scene from new view points. The user of this system controls the virtual camera's movement through the scene. Nevertheless, the virtual images are computed from static cameras. Such system can also compute a 3d reconstruction of the scene.
Our research focuses on real-time live video-based rendering, without preprocessing. We also focus on video-based rendering system using uncalibrated cameras (camera whose characteristics such as focal, orientation and position are unknown). Such techniques are less restrictive than standard techniques using calibrated cameras. Finally, we attempt to use only consumer hardware such as webcams in order to make this system accessible for every one.

Introduction

Le rendu à base de vidéos est un domaine de l'informatique à la frontière de la synthèse d'images et de la vision par ordinateur. Le but de cette discipline est de proposer des méthodes permettant de générer une nouvelle vue d'une scène à partir d'un ensemble d'images de cette scène. Le cas échéant, il est possible de générer une reconstruction 3d de la scène.

Ce domaine de recherche est relativement jeune mais très actif. Il est lui même composé de quatre sous-domaines pour lesquels les objectifs et les contraintes diffèrent. D'un côté, nous avons les méthodes dont le but est de générer des images les plus vraisemblables. Ces méthodes utilisent beaucoup de calculs et sont généralement très lentes, si bien qu'elles ne peuvent traiter que les images provenant d'une scène statique où tous les éléments sont immobiles. Ensuite viennent les méthodes plus rapides capables de générer des nouvelles vues d'une scène statique en temps réel. Viennent enfin les méthodes permettant de créer de nouvelles vues à partir d'une scène dynamique, il s'agit alors non plus de rendu à base d'images mais de rendu à base de vidéos. La plupart de ces méthodes effectuent un prétraitement sur les vidéos avant de générer de nouvelles vues en temps réel. Il s'agit de rendu en différé. Cependant les méthodes les plus rapides permettent de se passer de prétraitement et effectuent un rendu en direct. Notre méthode appartient à cette dernière catégorie.

Évidemment, la rapidité s'obtient souvent au détriment de la qualité, ce qui fait que ces quatre approches ne sont pas tout à fait concurrentes.

Objectifs et avancement :

Durant ma thèse, j'ai réalisé un dispositif de rendu à base de vidéos permettant de générer en direct et en temps réel de nouvelles images à partir de caméras calibrées, c'est-à-dire de caméras dont on connaît les paramètres (focale, position, orientation, etc.). La méthode que j'ai employé appartient à la famille des “plane-sweep” (balayage par plans). Cette méthode est décrite dans le chapitre suivant.

Le calibrage d'une caméra nécessite un ensemble de correspondances entre des points 3d de la scène et des pixels provenant des caméras. Ces correspondances sont parfois difficiles à obtenir et le fait d'utiliser des caméras calibrées impose que celles-ci soient fixes les unes par rapport aux autres durant le tournage, interdisant ainsi d'effectuer un zoom ou un recadrage bénéfique pour la qualité du rendu.

Notre objectif est de réaliser un dispositif utilisant des caméras non calibrées. Chaque rendu est alors précédé d'une étape de calibrage, autorisant ainsi les caméras à bouger et à zoomer. Nous souhaitons cependant conserver le rendu en temps réel et en direct. Le fait de pouvoir bouger les caméras présente un intérêt pour le rendu car il devient alors possible de placer les caméras dans une configuration optimale pour notre méthode de rendu, notamment en zoomant avec une des caméras sur une zone de la scène dont ont veut un maximum de détails. De plus, le dispositif de caméras devient beaucoup plus simple à utiliser car l'étape de calibrage n'est plus nécessaire.

Notre but est donc de mettre au point un dispositif de calibrage fonctionnant en temps réel qui soit suffisamment précis pour notre application.

Rendu à base de vidéos : méthode des « plane-sweep »

Le système de rendu à base de vidéos que j'ai mis au point durant ma thèse appartient à la famille des plane-sweep (balayage par plan) dont le mode de fonctionnement est expliqué dans ce chapitre.

La figure 1 représente la scène vue de dessus. Elle est filmée par deux caméras *cam1* et *cam2* supposées calibrées (on connaît donc leur position et leur orientation ainsi que leur caractéristiques techniques). La première étape consiste à placer la caméra virtuelle *camx* à partir de laquelle la nouvelle vue sera générée. Il faut alors définir un plan *near* et un plan *far* tels que tous les objets de la scène susceptibles d'être filmés soient situés entre ces deux plans. L'espace entre ces deux plans est alors discrétisé à l'aide de plans parallèles.

Considérons un point M de l'espace situé à la fois sur la surface d'un objet de la scène et sur un des plans. Ce point sera alors vu par les caméras avec la même couleur : la couleur de l'objet. Il n'en sera probablement pas de même si ce point M n'est pas situé sur la surface d'un objet. La méthode des plane sweep consiste donc à supposer que si un point d'un plan est vu par toutes les caméras avec la même couleur, c'est qu'il représente probablement la surface d'un objet de la scène.

Il suffit alors d'examiner pour chaque point de chaque plan la façon dont il est perçu par les caméras. On attribue à chaque point un score, celui-ci est bon si ce point est perçu par les caméras avec la même couleur et mauvais sinon. Une fois tous les points d'un plan traités, on projette les scores et les couleurs des points sur la caméra virtuelle. Pour chaque pixel de la caméra virtuelle, on ne met à jour une couleur que si son score est meilleur que celui de la couleur initiale.

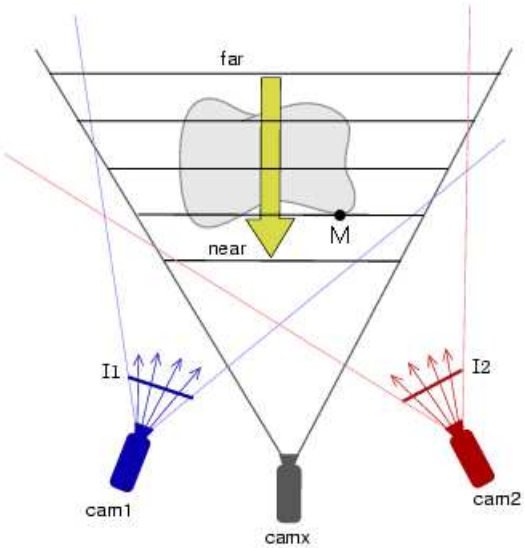

Figure 1 : *plane sweep, schémas*

Il se trouve que cette méthode est particulièrement bien adaptée à une implémentation sur la carte graphique d'un ordinateur en utilisant les *fragment shaders*. Le processeur de la carte graphique (GPU) ne sait faire que des opérations simples mais les fait plus vite que le processeur de l'ordinateur (CPU). C'est grâce à cette particularité que la méthode des plane-sweep génère des nouvelles images en temps réel.

Notons que contrairement aux méthodes concurrentes, la méthode des plane sweep n'utilise qu'un seul ordinateur et se contente de matériel “grand public” (webcams) ce qui la rend abordable pour tous. La figure 2 montre une nouvelle vue (320x240) générée à partir de 4 images. En subdivisant la scène avec 30 plans, cette méthode permet de générer 40 images par secondes avec une bonne carte graphique.

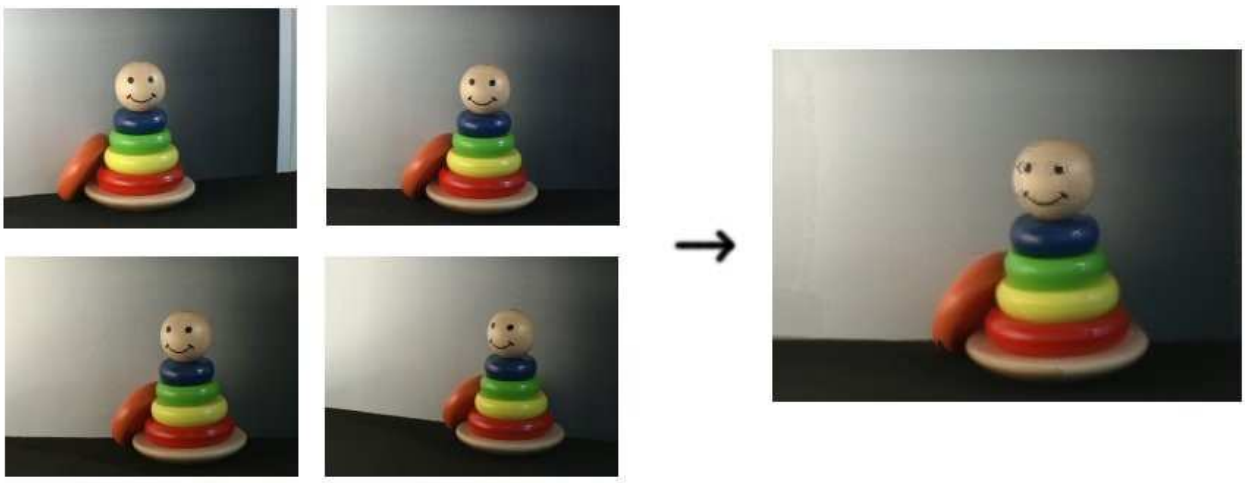

Figure 2 *: exemple de création d'une nouvelle vue en temps réel à partir de 4 caméras.*

Rendu à base de vidéos : caméras non calibrées

Il s'agit ici d'insérer une phase de calibrage avant la phase de rendu. Le calibrage d'une caméra consiste à déterminer ses caractéristiques comme sa focale, sa position, son orientation ou bien les particularités des capteurs CCD. Il existe divers méthodes de calibrage. Les plus robustes utilisent des correspondances entre des points 3d de la scène et leur projeté (coordonnée en pixel) sur la caméra à calibrer. Il est alors nécessaire de traiter chaque image afin d'y détecter et identifier les points 3d utilisés pour les correspondances. De plus, la position de ces points 3d doit être parfaitement connues. Ces deux contraintes sont difficilement conciliables avec les systèmes de rendu à base de vidéos.

Il existe une méthode de calibrage développée au sein de l'équipe du Professeur Hideo Saito permettant de calibrer une caméra simplement en identifiant sur les images des motifs carrés disposés dans la scène. Peu importe la taille et la disposition de ces carrés pourvu qu'ils ne soient pas placés sur un même plan. La déformation perspective de ces carrés filmés par les caméras suffit à extraire les informations nécessaires. Il est important de noter que la détection de carrés sur une image est relativement aisée. Il est probable que trois ou quatre carrés disposés dans la scène devraient suffire.

Un des avantages majeurs de cette méthode est qu'elle supporte tout à fait la manipulation des motifs carrés durant le tournage puisque leur position respective n'a pas d'incidence sur le calibrage. De plus, la taille des carrés n'a pas besoin d'être connue ce qui rend cette méthode très simple à utiliser.

En pratique, nous allons commencer nos tests en utilisant des marqueurs, faciles à suivre et à identifier sur une image. La figure 3 montre le genre d'images que nous comptons utiliser. On peut y distinguer deux marqueurs. Si cette méthode est concluante, nous envisageons d'utiliser des carrés

“naturels” comme par exemple une pochette de CD. Il est en effet fréquent de trouver des carrés dans les scènes filmées. La difficulté sera alors de les repérer et de les identifier sur chaque caméra.

Figure 3 : *marqueurs pour l'etape de calibrage*

Conclusion

Le rendu à base de vidéos est une discipline en plein essor. Nous essayons d'y contribuer en regroupant des méthodes de rendu à base de vidéos avec des méthodes de calibrage rapides. Le dispositif souhaité devrait pouvoir générer de nouvelles images d'une scène filmée avec plusieurs webcams et ce en temps réel et en direct. Ce dispositif devrait pouvoir fonctionner avec des caméras non calibrées.