



# Utilisation des longues focales lors des prises de vues stéréoscopiques

Sergi Pujades, Frédéric Devernay

► **To cite this version:**

Sergi Pujades, Frédéric Devernay. Utilisation des longues focales lors des prises de vues stéréoscopiques. Orasis, Congrès des jeunes chercheurs en vision par ordinateur, Jun 2013, Cluny, France. hal-00829395

**HAL Id: hal-00829395**

**<https://hal.archives-ouvertes.fr/hal-00829395>**

Submitted on 5 Jun 2013

**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.

# Utilisation des longues focales lors des prises de vues stéréoscopiques

Sergi Pujades  
sergi.pujades-rocamora@inria.fr

Inria Grenoble

Frédéric Devernay  
frederic.devernay@inria.fr

## Résumé

Nous analysons la problématique du changement de focale lors des prises de vues stéréoscopiques pour le cinéma ou la télévision 3D. Ces changements affectent deux éléments importants de l’affichage de la scène filmée : la rondeur des objets et la limite de divergence oculaire. Chaque élément détermine un entraxe des caméras. Selon la configuration, cet entraxe peut être différent et incompatible, d’autant plus que les focales sont longues. En pratique ce phénomène restreint l’utilisation des très longues focales lors des tournages. Afin de permettre son utilisation, nous proposons une solution basée sur l’utilisation de 3 caméras.

## Mots Clef

Cinéma stéréoscopique, Disparity Mapping, rondeur, divergence oculaire, distance focale.

## Abstract

We study the behavior of the rendered 3D scene when changing the focal length of a binocular acquisition system. Changing the focal length impacts two main features of the rendered scene : the roundness factor and the eye divergence. Each feature implies an interaxial distance between the cameras. In some cases those two distances are different and incompatible. The bigger the focal length, the bigger the incompatibility. In practice this leads to a restriction on the size of the focal length used for 3d acquisition. To allow the use of the big focal length lenses we propose a method based on 3 cameras.

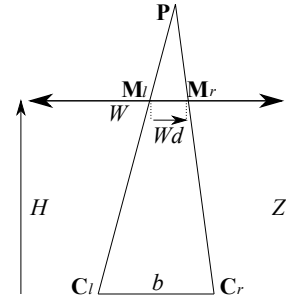
## Keywords

Stereoscopic cinema, disparity mapping, roundness factor, eye divergence, focal length.

## 1 Introduction

La création de contenu 3D est en plein essor. Néanmoins, certains usages du tournage de films 2D ne sont pas facilement transposables lors de l’acquisition 3D. Les longues focales en sont un exemple. Les longues focales sont souvent utilisées pour « se rapprocher de l’action » quand il est physiquement impossible d’approcher la caméra. Le sport ou les documentaires animaliers constituent des bons exemples où les longues focales sont souvent utilisées. Nous analysons les effets de son utilisation et proposons des pistes pour résoudre le problème.

## 2 De la captation à l’affichage



| Symbole    | Caméra  | Affichage            |
|------------|---|----------------------|
| $C_l, C_r$ | centre optique caméra                         | centre optique œil   |
| $P$        | point scène physique                          | point 3D perçu       |
| $M_l, M_r$ | points image de $P$                           | points écran         |
| $b$        | entraxe caméras                               | interoculaire humain |
| $H$        | distance de convergence                       | distance à l’écran   |
| $W$        | largeur plan convergence                      | taille écran         |
| $Z$        | profondeur réelle                             | profondeur perçue    |
| $d$        | disparité gauche-droite (en fraction de $W$ ) |                      |

FIG. 1 – Paramètres décrivant la géométrie du tournage et celle de l’affichage.

Nous pouvons formaliser le processus de capture d’une scène 3D et son affichage dans la salle de projection comme une transformation de l’espace 3D  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  qui dépend des paramètres d’acquisition et d’affichage. Ces paramètres sont détaillés sur la Fig. 1 [1]. Dans notre étude, nous nous intéressons aux déformations de la profondeur. La profondeur perçue  $Z'$ , en fonction des paramètres de capture ( $b, H, W$  et  $Z$ ) et d’affichage ( $b', H'$ , et  $W'$ ) est :

$$Z' = \frac{H'}{1 - \frac{W'}{b'} \left( \frac{b}{W} \frac{Z-H}{Z} \right)}. \quad (1)$$

Deux éléments importants lors de l’affichage de la scène sont d’éviter la divergence oculaire et d’assurer la rondeur souhaité du sujet filmé. Nous étudions ensuite ses relations avec les paramètres  $H$  et  $W$ .

### 2.1 Rondeur d’un élément

La rondeur d’un élément  $\rho$  est définie comme le rapport entre la variation de profondeur apparente  $\frac{\partial Z'}{\partial Z}$  et sa variation de largeur apparente  $\frac{\partial X'}{\partial X}$  [1]. Si on se place dans le

plan de l'écran, l'équation se simplifie pour  $\rho_{ecran}$  :

$$\rho = \frac{\partial Z'}{\partial Z} / \frac{\partial X'}{\partial X}, \rho_{ecran} = \frac{b}{H} \frac{H'}{b'}. \quad (2)$$

Si on veut garder une rondeur constante sur le sujet,  $b$  doit être proportionnel à  $H$  :

$$b_{rond} = \rho \frac{b'}{H'} H = kH. \quad (3)$$

## 2.2 Divergence oculaire

La divergence oculaire arrive si et seulement si [1] :

$$\frac{b}{W} < \frac{b'}{W'}, \text{ donc } b_{div} < \frac{b'}{W'} W. \quad (4)$$

## 2.3 Problématique du changement de focale

Lors d'un changement de focale, les paramètres d'acquisition sont modifiés. Si l'objet reste à la même distance ( $H \rightarrow$ ), un incrément de focale fera diminuer  $W$  : ( $H \rightarrow, W \downarrow$ ). Si l'objet s'éloigne ( $H \uparrow$ ), on peut augmenter la focale pour le garder avec la même taille sur l'écran ( $W \rightarrow$ ) : ( $H \uparrow, W \rightarrow$ ). Ou si l'objet s'approche ( $H \downarrow, W \rightarrow$ ).

Les contraintes (3) et (4) dépendent des paramètres  $H$  et  $W$  respectivement. Et la modification de la focale donne des  $b_{rond}$  et  $b_{div}$  incompatibles :

- Si ( $H \uparrow, W \rightarrow$ )  $\Rightarrow$  ( $b_{rond} \uparrow, b_{div} \rightarrow$ )
- Si ( $H \rightarrow, W \downarrow$ )  $\Rightarrow$  ( $b_{rond} \rightarrow, b_{div} \downarrow$ )

Le rapport entre  $b_{rond}$  et  $b_{div}$  est proportionnel au rapport entre  $H$  et  $W$ . Dans le cas des très longues focales ce dernier peut facilement atteindre 50 : 1, impliquant une forte incompatibilité entre les entraxes théoriques.

## 2.4 Travaux antérieurs

Bien que la problématique des longues focales n'ait pas été traitée explicitement, dans la littérature plusieurs auteurs ont traité le problème du *disparity mapping* [2, 3, 4]. Il s'agit de modifier la disparité des éléments de la scène pour éviter la divergence oculaire, leur donner la rondeur désirée ou encore d'autres effets artistiques. Dans ces méthodes, on retrouve une fonction  $\phi(d) : \mathbb{R} \rightarrow \mathbb{R}$  qui transforme la disparité originale d'une paire stéréoscopique en celle souhaitée.

Devernay et al. [1] traitent le problème de changement de taille d'écran et proposent une méthode de rendu basée sur l'interpolation de points de vues. Lang et al. [3] proposent une méthode de *warp d'image* basée sur des indices de *saillance*[6], afin de respecter les parties de l'image qui vont attirer l'attention visuelle. Tan, Lau et al. [4] s'inspirent fortement de [3] mais proposent d'utiliser des éléments haut niveau (droites et plans) plutôt que la saillance. Ils obtiennent des très bons résultats visuels. Dans cet article, ils appliquent une transformation de disparité avec un rapport 2 : 1, qu'ils considèrent déjà comme une *grande déformation*. Dans toutes les méthodes proposées,  $\phi(d)$  n'introduit jamais des déformations de l'image équivalentes à l'utilisation des longues focales.

Dans le cas d'images de synthèse, Neuman [5] propose une technique dite *multi-rigging*. Cette méthode utilise plusieurs configurations de caméras pour faire plusieurs rendus stéréoscopiques de chaque élément de la scène, puis les mélanger. Il a donc la possibilité de faire un rendu avec  $b_{div}$  et un avec  $b_{rond}$ , puis composer le résultat en une seule image finale en utilisant l'information de profondeur de sa scène.

## 2.5 Approche proposée

Pour la prise de vues réelle nous proposons une méthode analogue à [5] utilisant trois caméras : la première et la deuxième avec un entraxe  $b_{div}$  (eq. 4), et la première et la troisième avec un entraxe  $b_{rond}$  (eq. 3). À partir de ces trois images, nous cherchons à composer l'image qui respectera à la fois la règle des infinis et la rondeur du sujet.

Quand on veut générer une image à partir d'autres images il faut calculer la transformation de chaque image vers l'image finale, puis mélanger ces images transformées. Les fonctions de transformation géométriques requièrent le plus souvent une connaissance, même approximative, de la géométrie de la scène. Nous aimerons étudier comment mélanger les images en s'inspirant sur des approches provenant de deux problématiques bien étudiées : les méthodes de rendu basées image [7] et les méthodes variationnelles du domaine de la super-résolution [8].

Buheler et al.[9] proposent des directives pour la synthèse de nouveaux points de vue à partir d'un ensemble d'images. Dans son cas, la géométrie est toujours fixe. Nous souhaitons proposer un cadre unifié qui englobe [9] et qui prenne en compte les possibles variations de la géométrie, en nous basant sur des mesures des déformations des images originales.

## Références

- [1] Devernay and Beardsley, *Image and Geometry Processing for 3-D Cinematography*, Springer-Verlag 2010.
- [2] Devernay et al., *Adapting stereoscopic movies to the viewing conditions using depth-preserving and artifact-free novel view synthesis*, SD&A 2011.
- [3] Lang et al., *Nonlinear Disparity Mapping*, SIGGRAPH 2010.
- [4] Tan et al., *Depth Mapping for stereoscopic videos*, IJCV 2012.
- [5] P. Neuman, *Bolt 3D : a case study*, SD&A 2010.
- [6] Guo et al., *Spatio-temporal saliency detection using phase spectrum of quaternion Fourier transform*, CVPR 2008.
- [7] Shum et al., *Image-Based Rendering*, Springer-Verlag, 2007.
- [8] Kang et al., *SuperResolution Image Reconstruction : a technical Overview*, IEEE Signal Processing Magazine, 2003.
- [9] Buheler et al. *Unstructured Lumigraph Rendering*, SIGGRAPH 2001.