

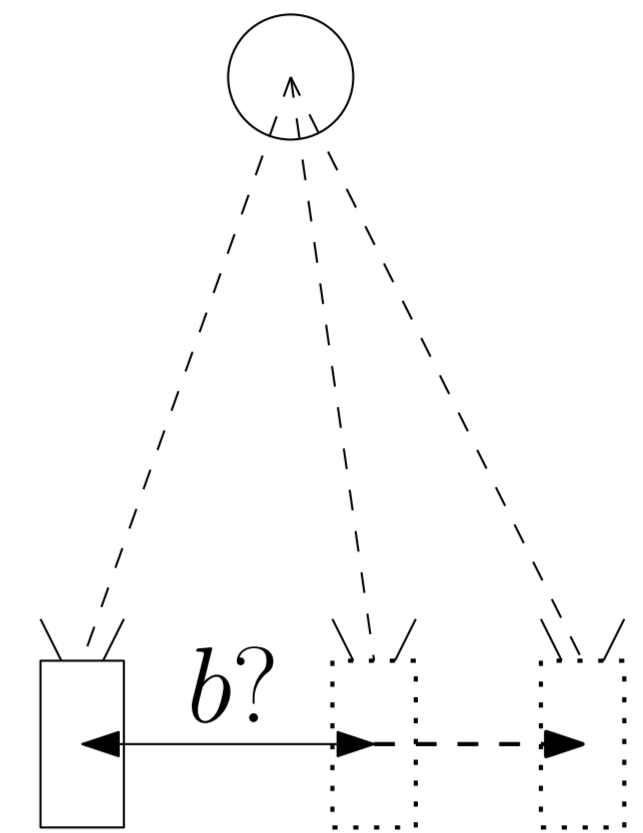
Contexte : Cinéma ou TV 3D, captation et affichage

Question : quelle séparation de caméras utiliser?



Étant données :

- une scène à filmer
  - une salle pour projeter
- Quelle séparation choisir?



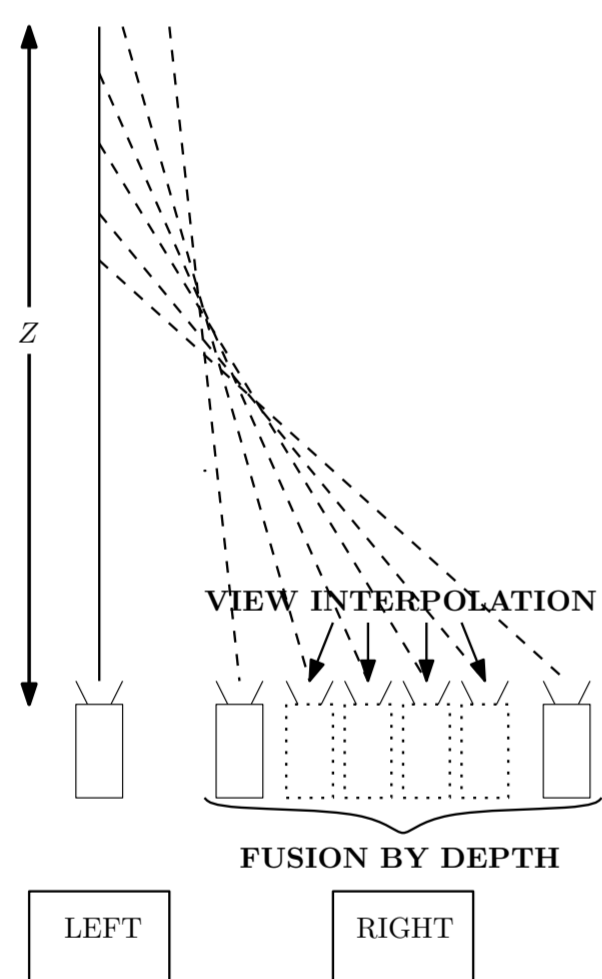
Petite séparation : éviter la divergence oculaire ( $b \approx 19.5 \text{ cm}$ )

Grande séparation : préserver la rondeur du sujet ( $b \approx 65 \text{ cm}$ )



Méthode envisagée

Résultat souhaité



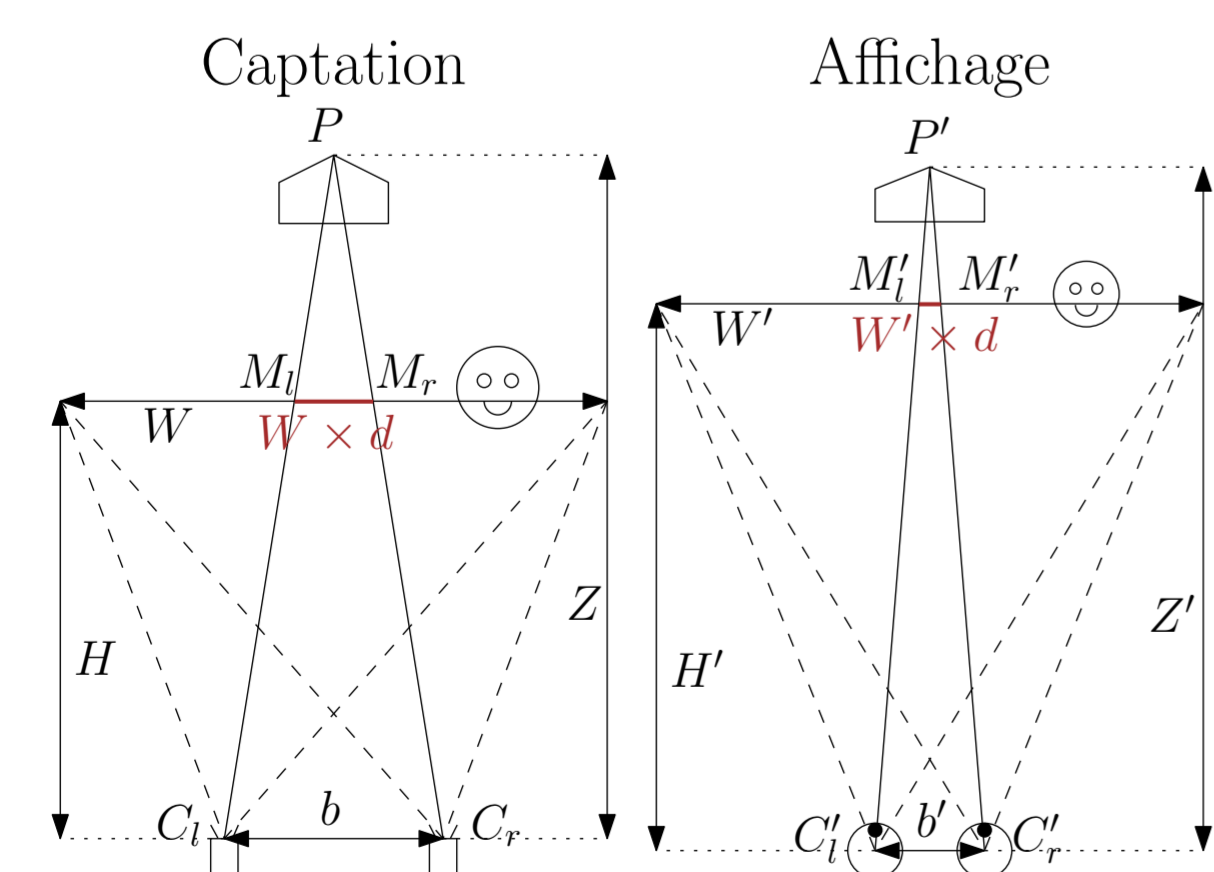
- Génération de points de vue intermédiaires
- *Compositing* de points de vue selon la profondeur



De la captation à l'affichage

Formalisation des contraintes

Impact du changement de focale



Profondeur perçue :

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

Rondeur élément :  $\rho = \frac{\partial Z'}{\partial X'}$

Rondeur élément sur l'écran :

$$\rho_{\text{ecran}} = \left. \frac{\partial Z'}{\partial X'} \right|_{Z=H} = \frac{b}{H} \frac{H'}{W'}$$

Séparation rondeur :

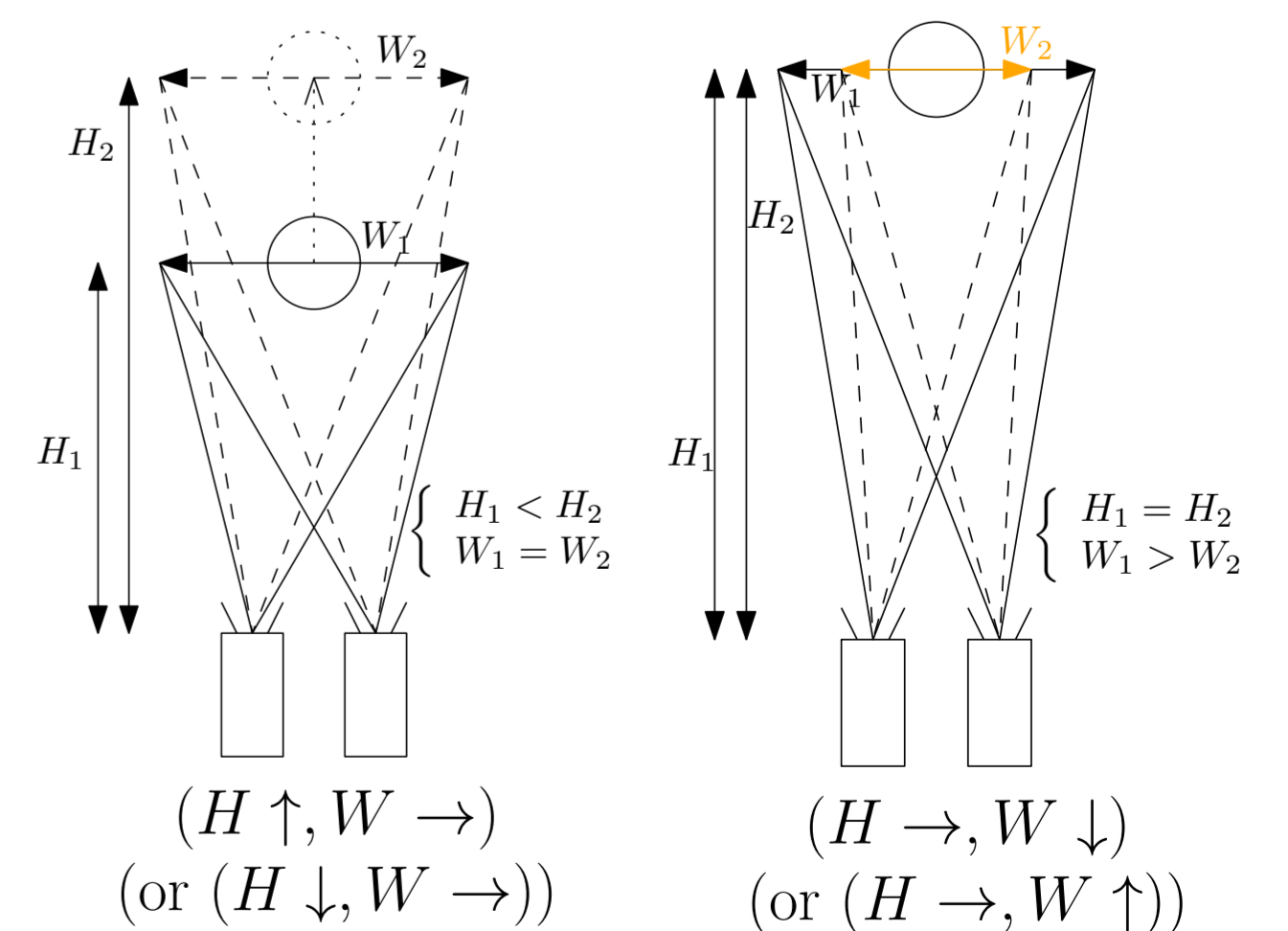
$$b_{\text{rond}} = \rho \frac{b'}{H'} H$$

Limite de divergence :

$$\frac{b}{W} < \frac{b'}{W'}$$

Séparation maximale sans divergence :

$$b_{\text{div}} = \frac{b'}{W'} W$$



Incohérence des séparations :

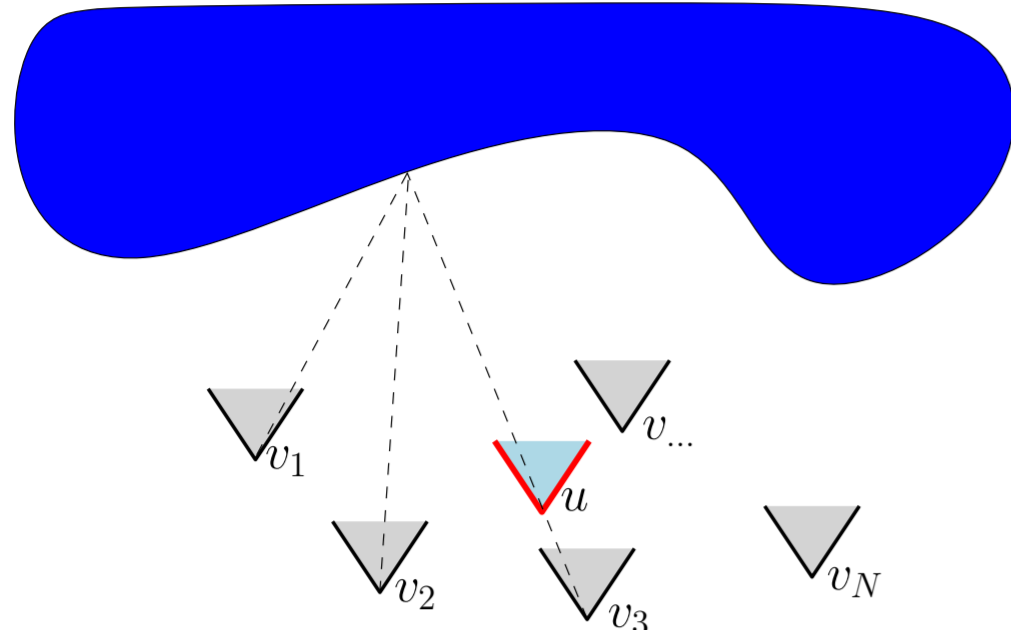
- Si  $(H \uparrow, W \rightarrow) \Rightarrow (b_{\text{rond}} \uparrow, b_{\text{div}} \rightarrow)$
  - Si  $(H \rightarrow, W \downarrow) \Rightarrow (b_{\text{rond}} \rightarrow, b_{\text{div}} \downarrow)$
- $$\frac{b_{\text{rond}}}{b_{\text{div}}} \sim \frac{H}{W} \sim \text{focal}$$

Symbole	Captation	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 $W$ )	



## Problème [5]

- Entrée : ensemble d'images  $v_i, i \in 1, \dots, N$
- Générer une image  $u$



## Objectifs [2]

- Utilisation de la géométrie disponible
- Configuration d'entrée non-structurée
- Cohérence épipolaire
- Distance angulaire minimale
- Continuité
- Sensibilité à la résolution
- Cohérence de rayons équivalents
- Temps réel

## Pré Calculs

- Calibration projective / métrique
- Géométrie approximative : disparité  $d_i$ 
  - dense avec confiance  $\sigma_{d_i}$
  - ou clairsemée (points clé)
- Autres informations
  - carte de saillance  $S_i$  [4]
  - structures (droites, ...)

## Méthodes directes

### Étapes

- Calcul des transformations  
 $\tau_i : v_i \rightarrow u$
- Mélange des images transformées

$$\tilde{u} = \sum_{i=1}^N \sum_{j=1}^{L \times H} w_i(j) v_i \circ \tau_i^{-1}(j)$$

avec

$$\forall j, \sum_{i=1}^N w_i(j) = 1$$

### Calcul des transformations $\tau_i : v_i \rightarrow u$

- Warp avec la géométrie dense [3]
- Warp de maillage avec points clé + saillance [6]
- Warp de maillage avec points clé + structures haut niveau [7]



### Calcul des poids de mélange $w_i(j)$

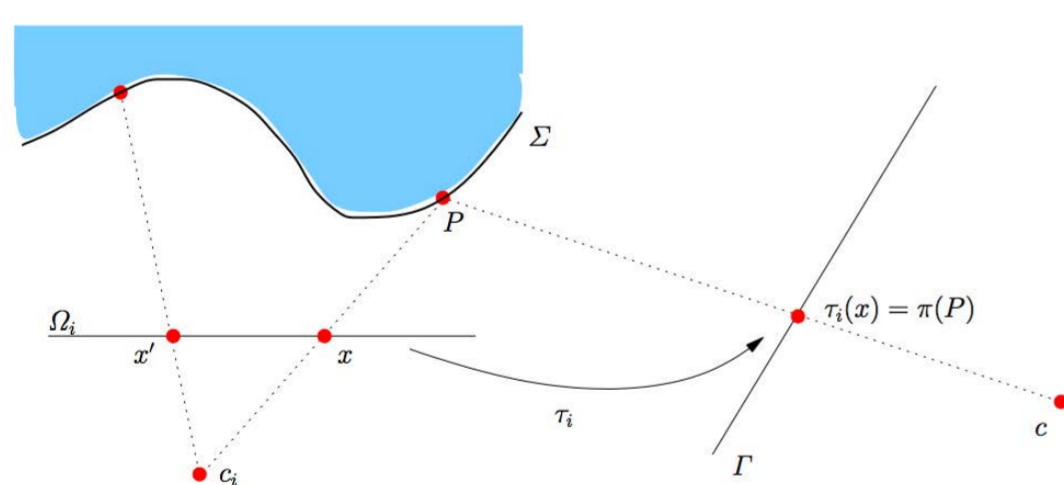
- Gestion des occlusions
- Poids heuristiques [2]
  - Cohérence épipolaire
  - Distance angulaire minimale
  - Continuité
  - Sensibilité à la résolution
  - Cohérence de rayons équivalents

## Méthodes variationnelles

### Étapes

- Modèle génératif
- Énergie issue du Maximum à Posteriori (MAP)
- Équations d'Euler-Lagrange
- Optimisation

### Modèle génératif existant [8]

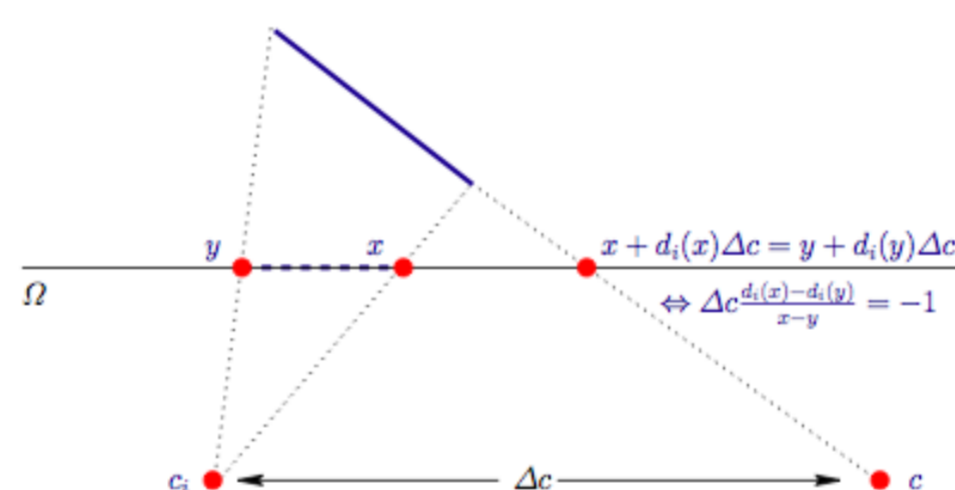


- Énergie issue du MAP :  
( $\sigma_s^2 :=$  bruit du capteur, gaussien, additif)

$$E(u) = \int_{\Gamma} |Du| + \frac{\sum_{i=1}^N \frac{1}{2} \int_{\Omega_i} m_i (b * (u \circ \tau_i) - v_i)^2 dx}{\sigma_s^2}$$

- Poids de mélange si  $\tau_i(x) = x + d_i(x) \cdot (c - c_i)$  :

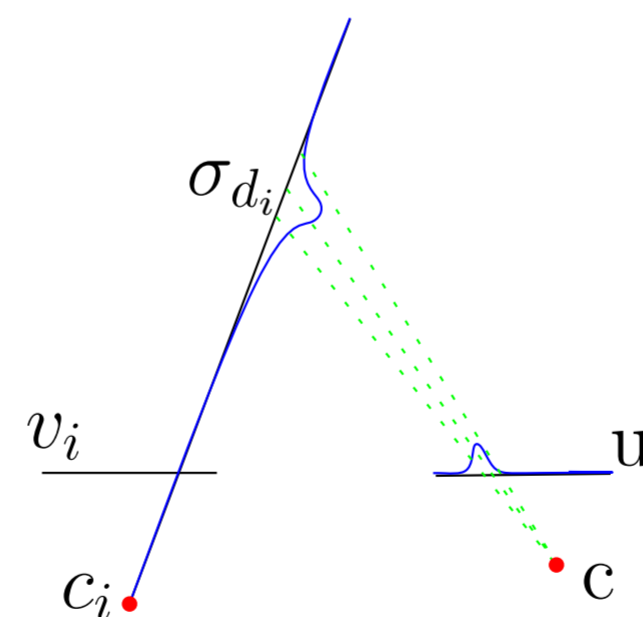
$$w_i(j) = \left| (1 + \nabla d_i \cdot (c - c_i)) \sigma_s^2 \right|^{-1}$$



### Modèle génératif proposé

- Nous considérons le bruit ajouté par l'imprécision de l'estimation de la géométrie :

$$\sigma_i^2 = \sigma_s^2 + \sigma_{g_i}^2$$



- Énergie issue du MAP

$$E(u) = \int_{\Gamma} |Du| + \frac{\sum_{i=1}^N \frac{1}{2} \int_{\Omega_i} m_i (b * (u \circ \tau_i) - v_i)^2 dx}{\sigma_s^2 + \sigma_{d_i}^2 (b * (\nabla u \cdot (c - c_i)))^2}$$

- Poids de mélange

$$w_i(j) = \left| (1 + \nabla d_i \cdot (c - c_i)) \left( \sigma_s^2 + \sigma_{d_i}^2 (b * (\nabla u \cdot (c - c_i)))^2 \right) \right|^{-1}$$

### Optimisation de la fonctionnelle

- Énergie Convexe  $\rightarrow$  FISTA [1]
- Méthode des moindres carrés itératif pondérés :
  - Évaluer poids de mélange avec solution courante  $\rightarrow$  FISTA

## References

- [1] Beck and Teboulle, Fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM 2009
- [2] Buehler et al. Unstructured Lumigraph Rendering, SIGGRAPH 2001.
- [3] Devernay et al., Adapting stereoscopic movies to the viewing conditions using depth-preserving and artifact-free novel view synthesis, SDA 2011.

- [4] Guo et al., Spatio-temporal saliency detection using phase spectrum of quaternion Fourier transform, CVPR 2008.
- [5] Shum et al., Image-Based Rendering, Springer-Verlag, 2007.
- [6] Lang et al., Nonlinear Disparity Mapping, SIGGRAPH 2010.
- [7] Tan et al., Depth Mapping for stereoscopic videos, IJCV 2012.
- [8] S. Wanner and B. Goldluecke, Spatial and Angular Variational Super-Resolution of 4D Light Fields, ECCV 2012.