Vision par ordinateur

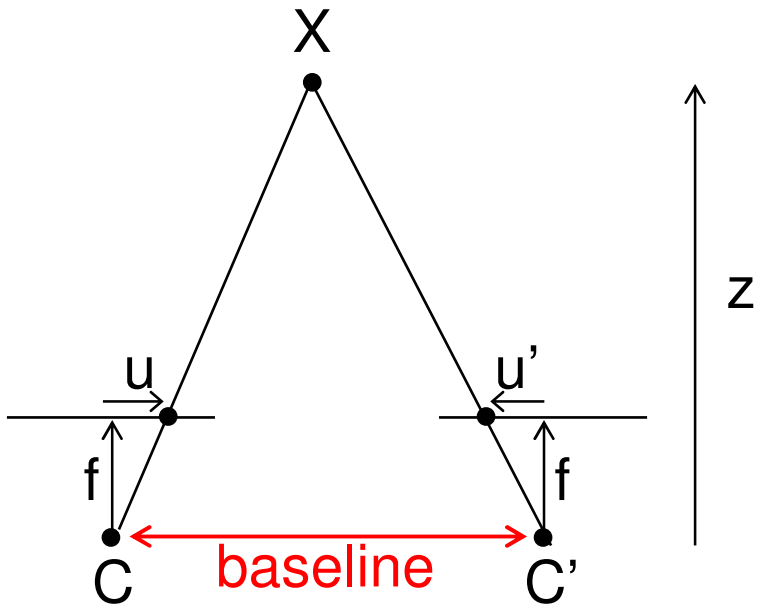
Stéréoscopie
multi-vues

Frédéric Devernay
d'après le cours de Steve Seitz

Reconstruction stéréo

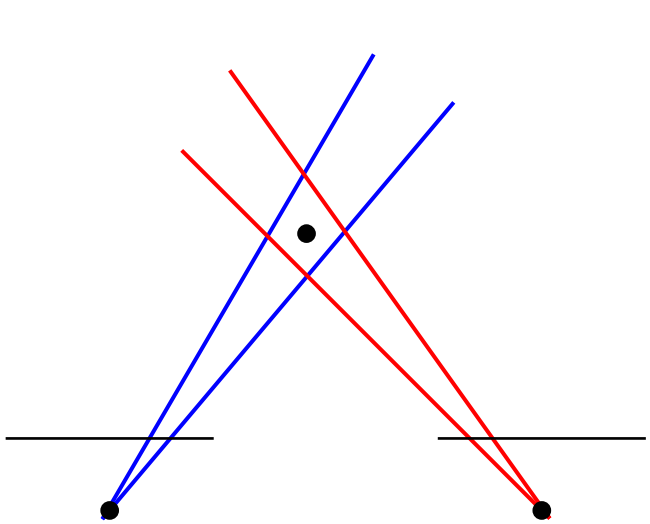
Étapes

- Calibrer les cameras
- Rectifier les images
- Calculer la disparité
- Estimer la profondeur

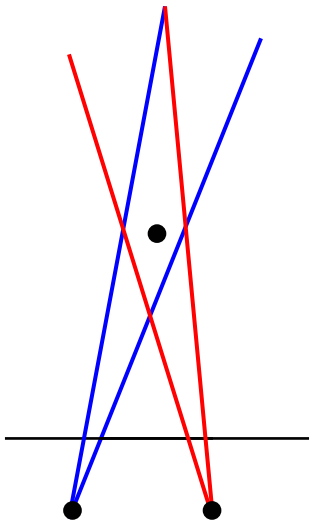


$$disparity = u - u' = \frac{baseline * f}{z}$$

Choix de la ligne de base (baseline)



Grande baseline



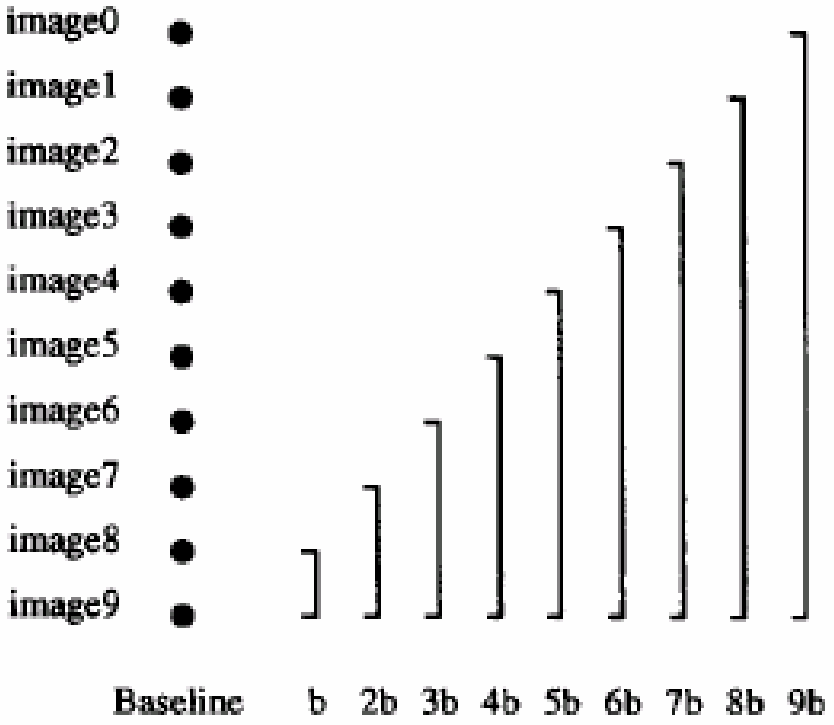
Petite baseline

Quelle est la baseline optimale ?

- Trop petite : grande erreur sur la profondeur
- Trop grande : mise en correspondance plus difficile

Effet de la baseline sur l'estimation de profondeur

Figure 2: An example scene. The grid pattern in the background has ambiguity of matching.



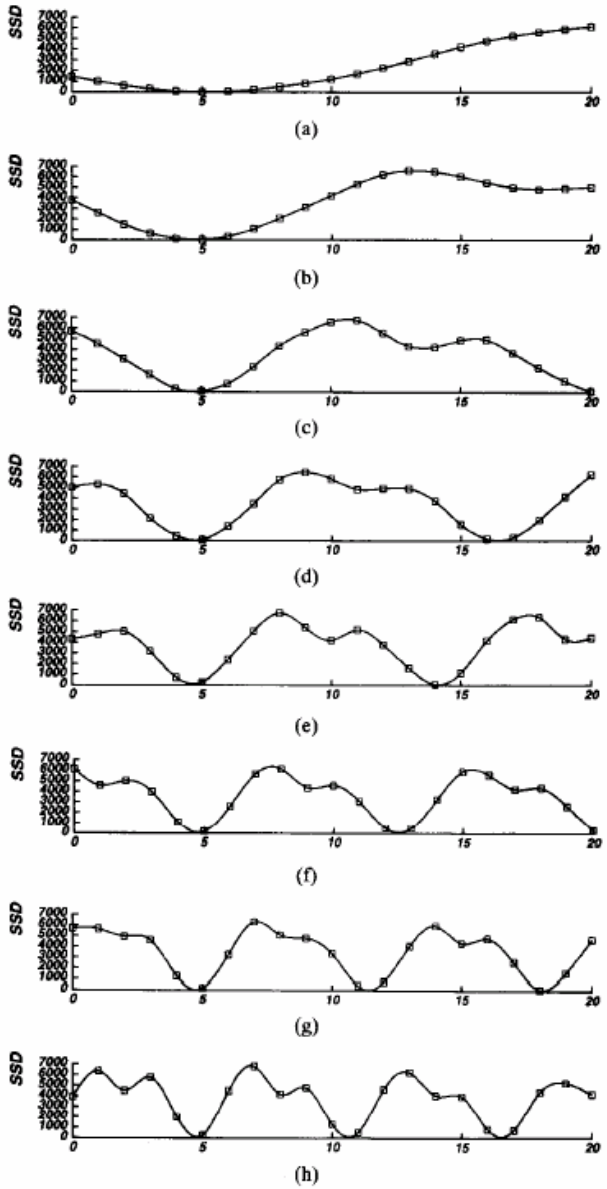


Fig. 5. SSD values versus inverse distance: (a) $B = b$; (b) $B = 2b$; (c) $B = 3b$; (d) $B = 4b$; (e) $B = 5b$; (f) $B = 6b$; (g) $B = 7b$; (h) $B = 8b$. The horizontal axis is normalized such that $8bF = 1$.

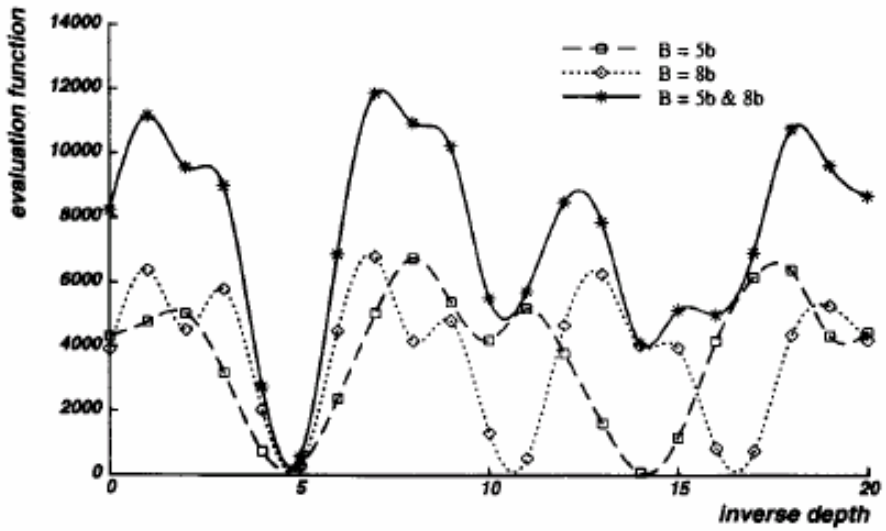


Fig. 6. Combining two stereo pairs with different baselines.

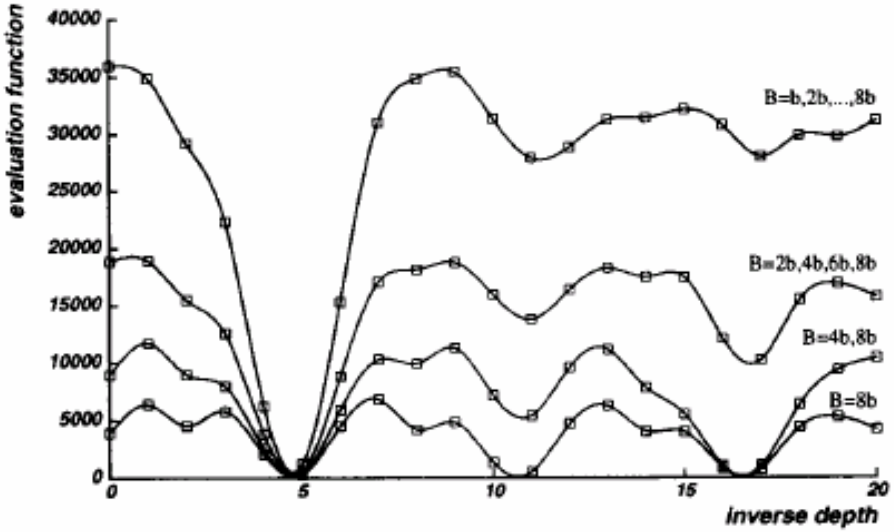


Fig. 7. Combining multiple baseline stereo pairs.

Stéréoscopie multi-baseline

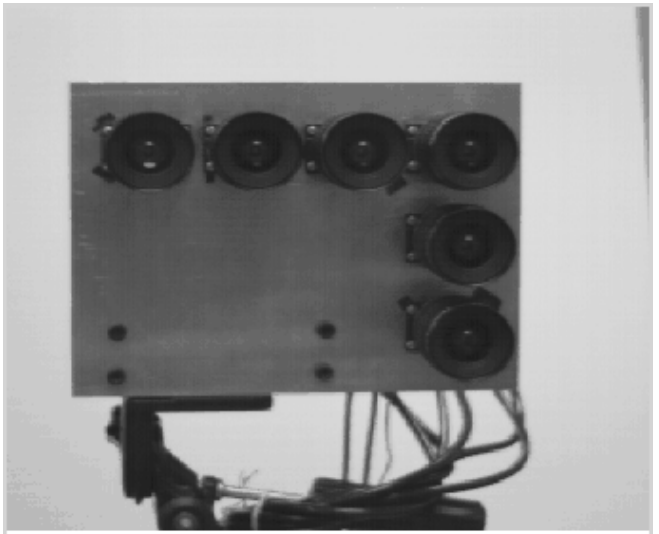
Approche de base

- Choisir une image de référence
- Utiliser n'importe quel algo de stéréo MAIS
  > remplacer le critère sur 2 vues par le critère sur toutes les baselines

Limitations

- On doit choisir une image de référence (asymétrique)
- Les points doivent être visibles sur toutes les images

Système vidéo temps-réel multi-baseline de CMU



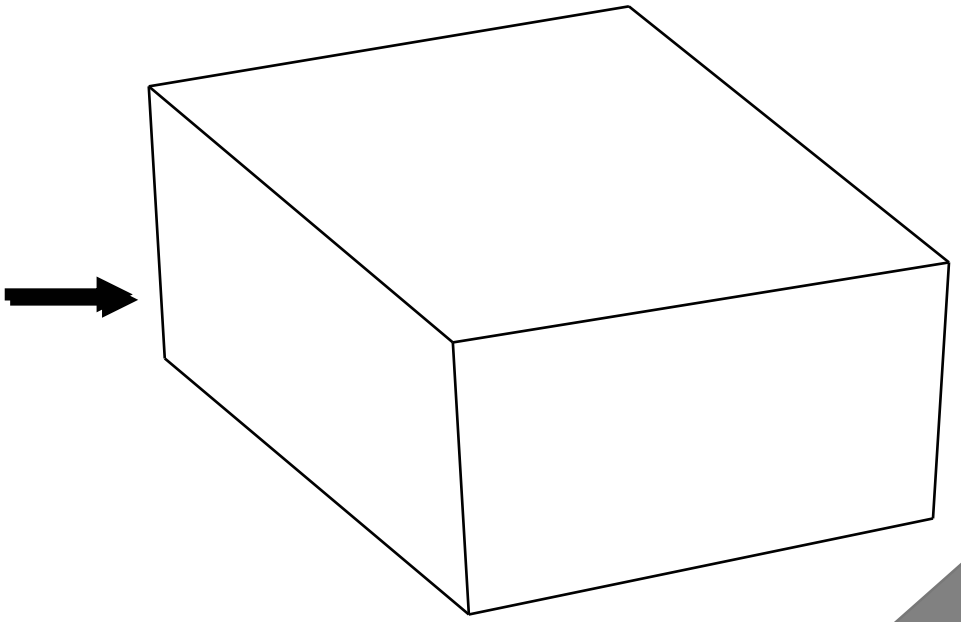
(a)

(b)

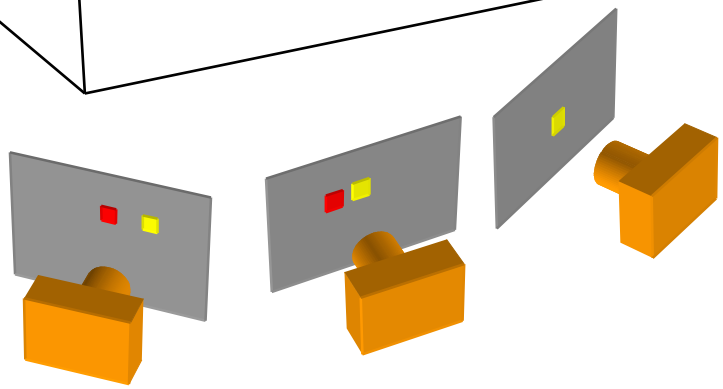
Figure 7: The CMU Video-Rate Stereo Machine Prototype System: (a) camera head; (b) processor boards

Stéréo volumétrique

Volume de
la scène **V**



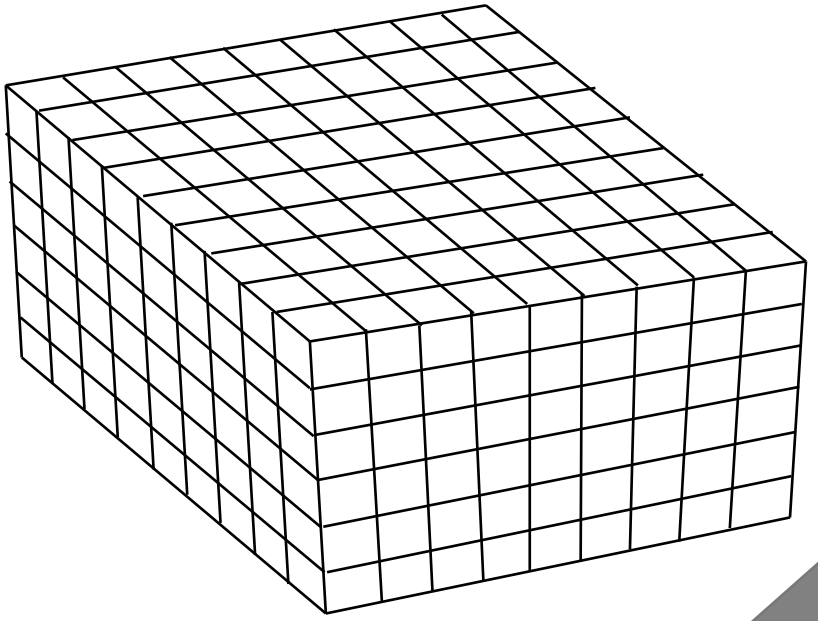
Images d'entrée
(Calibrées)



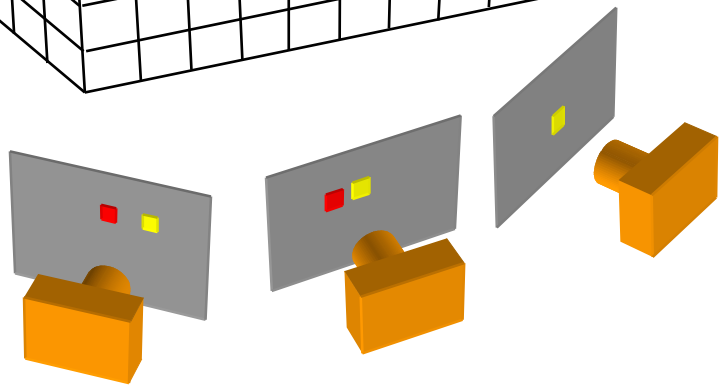
But : Déterminer la radiance des points de V

Formulation discrète : coloriage de voxels

Volume de la scène discrétisé



Images d'entrée (Calibrées)



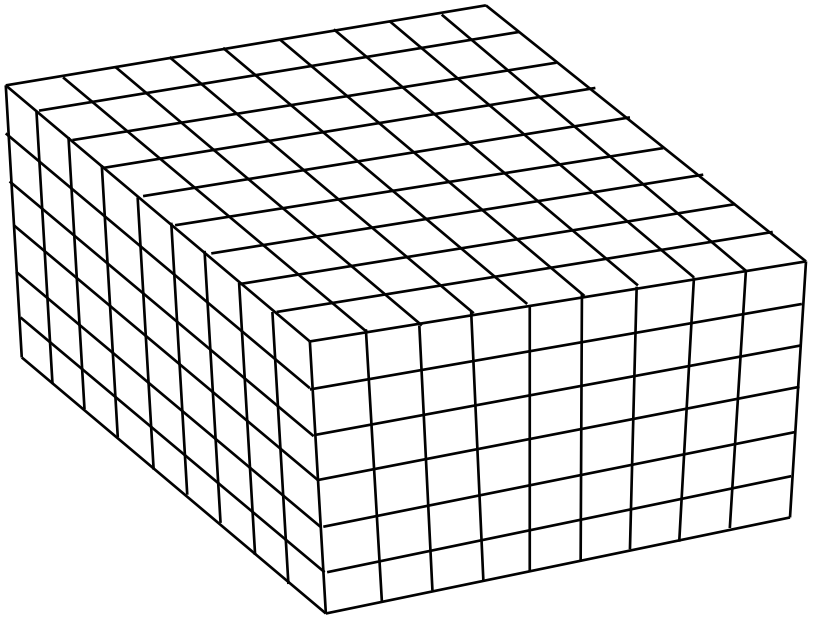
But : Attribuer des valeurs RGBA *photo-cohérentes* avec les images aux voxels de V

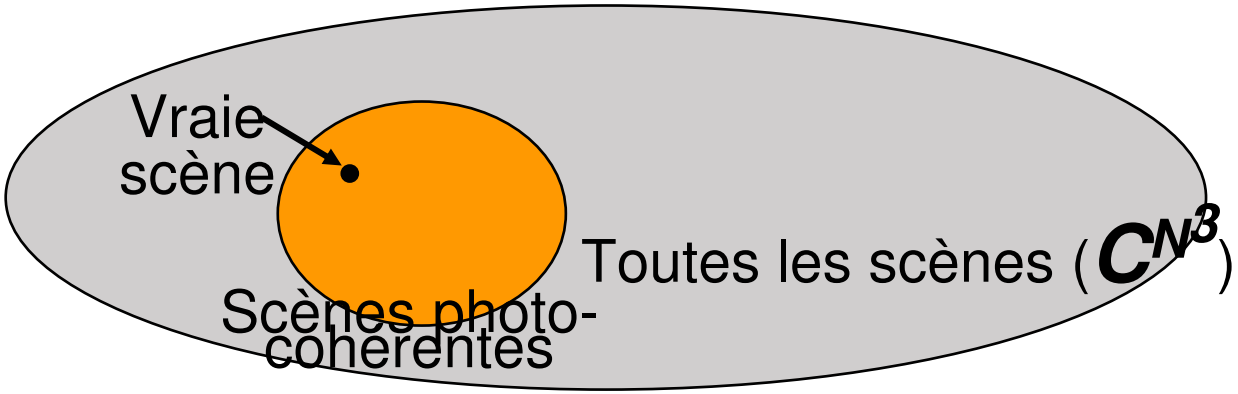
Complexité et calculabilité

Volume de la scène discrétisé

N^3 voxels

C couleurs





Problèmes

Questions théoriques

- Identifier la classe de *toutes* les scènes photo-cohérentes

Questions pratiques

- Comment calculer les modèles photo-cohérents ?

Solutions au coloriage de voxels

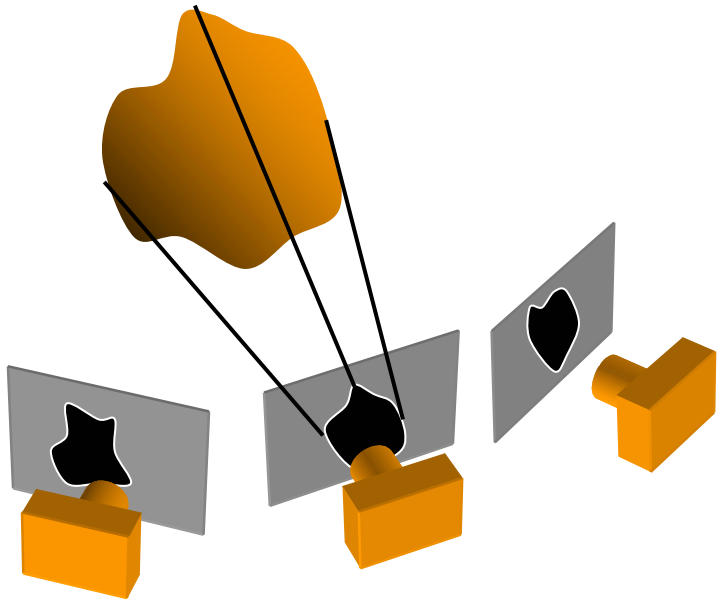
1. C=2 (silhouettes)
   - Intersection de volumes [Martin 81, Szeliski 93]

2. C non contraint, contraintes sur les points de vue
   - Algorithme de coloriage de voxels [Seitz & Dyer 97]

3. Cas général
   - Space carving (creusage de l'espace) [Kutulakos & Seitz 98]

Reconstruction à partir de silhouettes (C = 2)

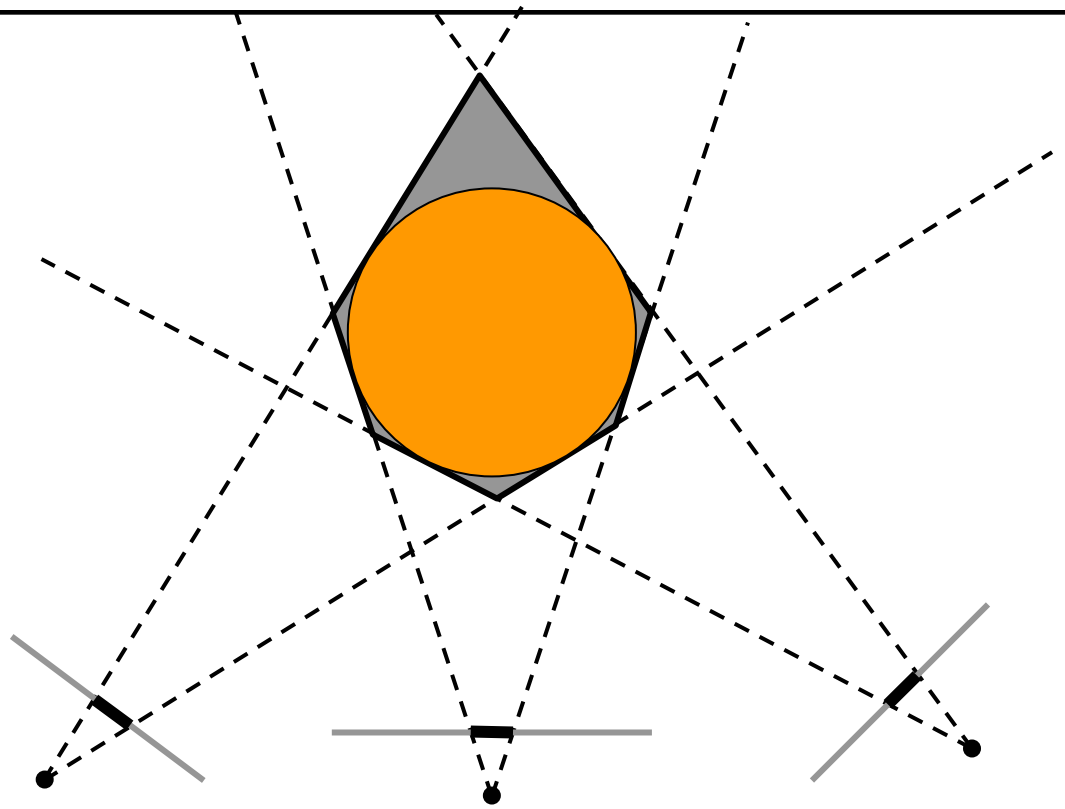
Images binaires →



Approche :

- *Rétroprojeter* chaque silhouette
- Intersecter les volumes rétroprojetés

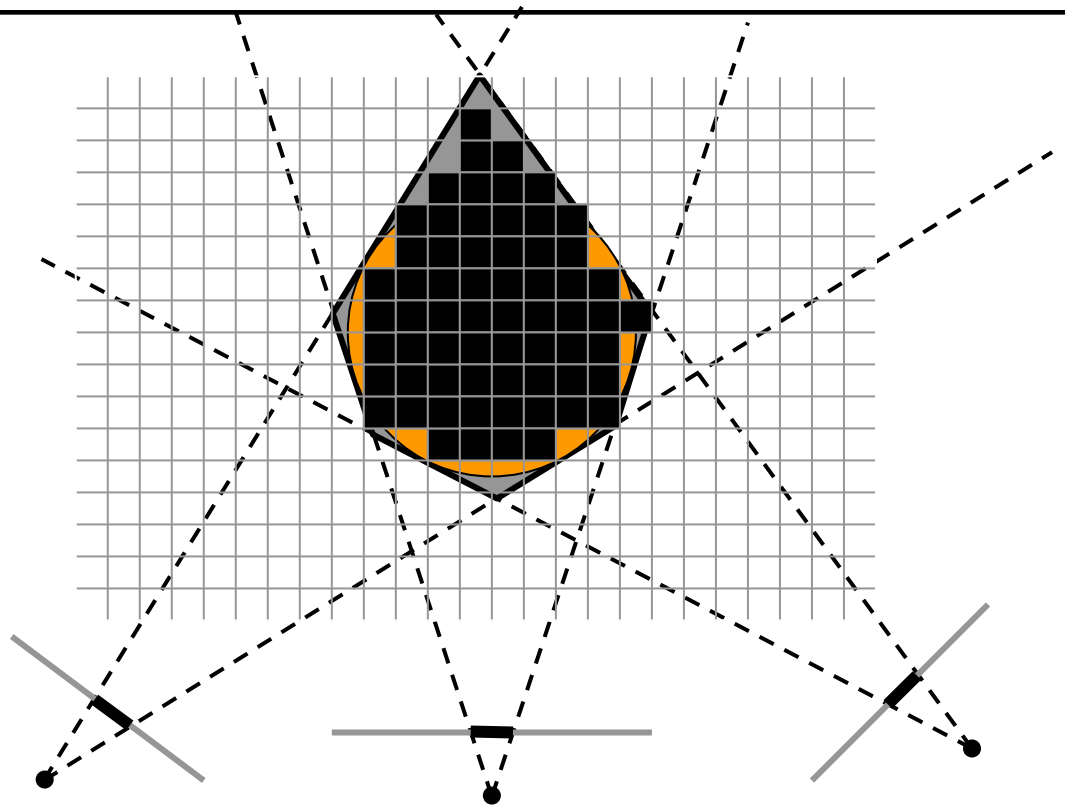
Intersection de volumes



La reconstruction contient la vraie scène

- mais n'est pas la même en général
- Au mieux on obtient *l'enveloppe convexe* ou plutot le complémentaire de toutes les droites qui n'intersectent pas S

Algorithme pour l'intersection de volumes



Colorier le voxel en noir s'il est dans la silhouette dans toutes les images

- $O(MN^3)$, pour M images, N^3 voxels
- Par besoin de tester les 2^{N^3} scènes possibles !

Propriétés de l'intersection de volumes

Pour

- Facile à implanter, rapide
- Accéléré par l'utilisation d'octrees [Szeliski 1993]

Contre

- Pas de concavités
- La reconstruction n'est pas photo-cohérente
- Nécessite l'identification des silhouettes

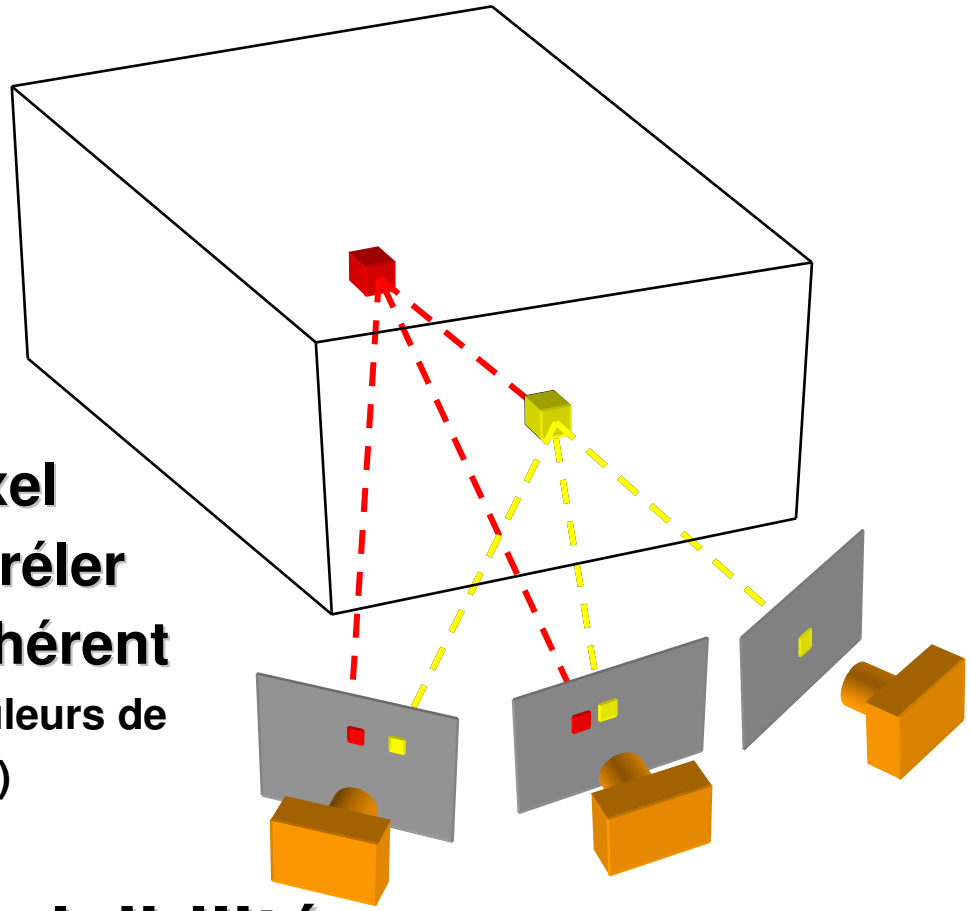
Solutions au coloriage de voxels

1. C=2 (silhouettes)
   - Intersection de volumes [Martin 81, Szeliski 93]

2. C non contraint, contraintes sur les points de vue
   - Algorithme de coloriage de voxels [Seitz & Dyer 97]

3. Cas général
   - Space carving (creusage de l'espace) [Kutulakos & Seitz 98]

Approche par coloriage de voxel



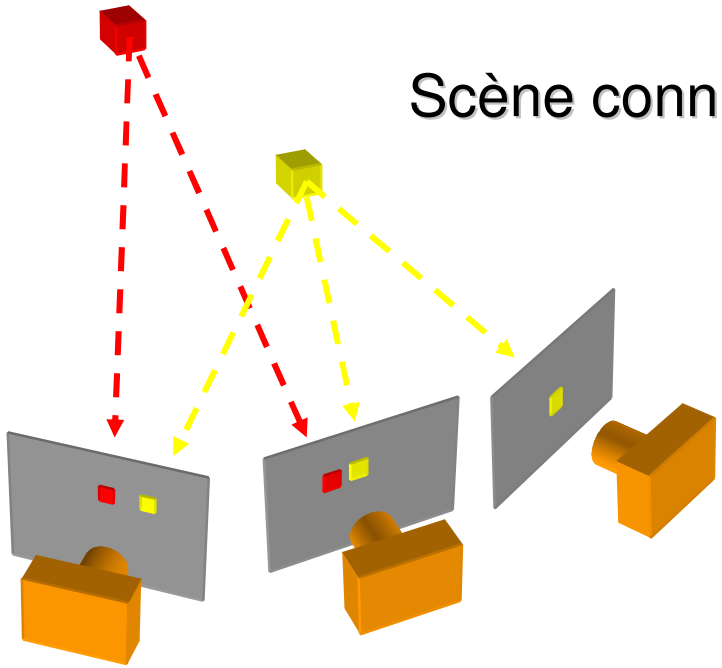
1. **Choisir un voxel**
2. **Projeter et corréler**
3. **Colorier si cohérent**
   (écart type des couleurs de pixel sous un seuil)

Problème de visibilité :

dans quelles images chaque voxel est-il visible?

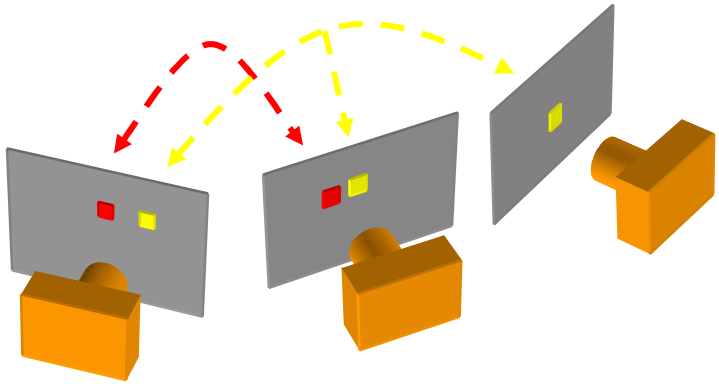
Le problème de visibilité globale

Quels points sont visibles dans quelles images ?



Scène connue

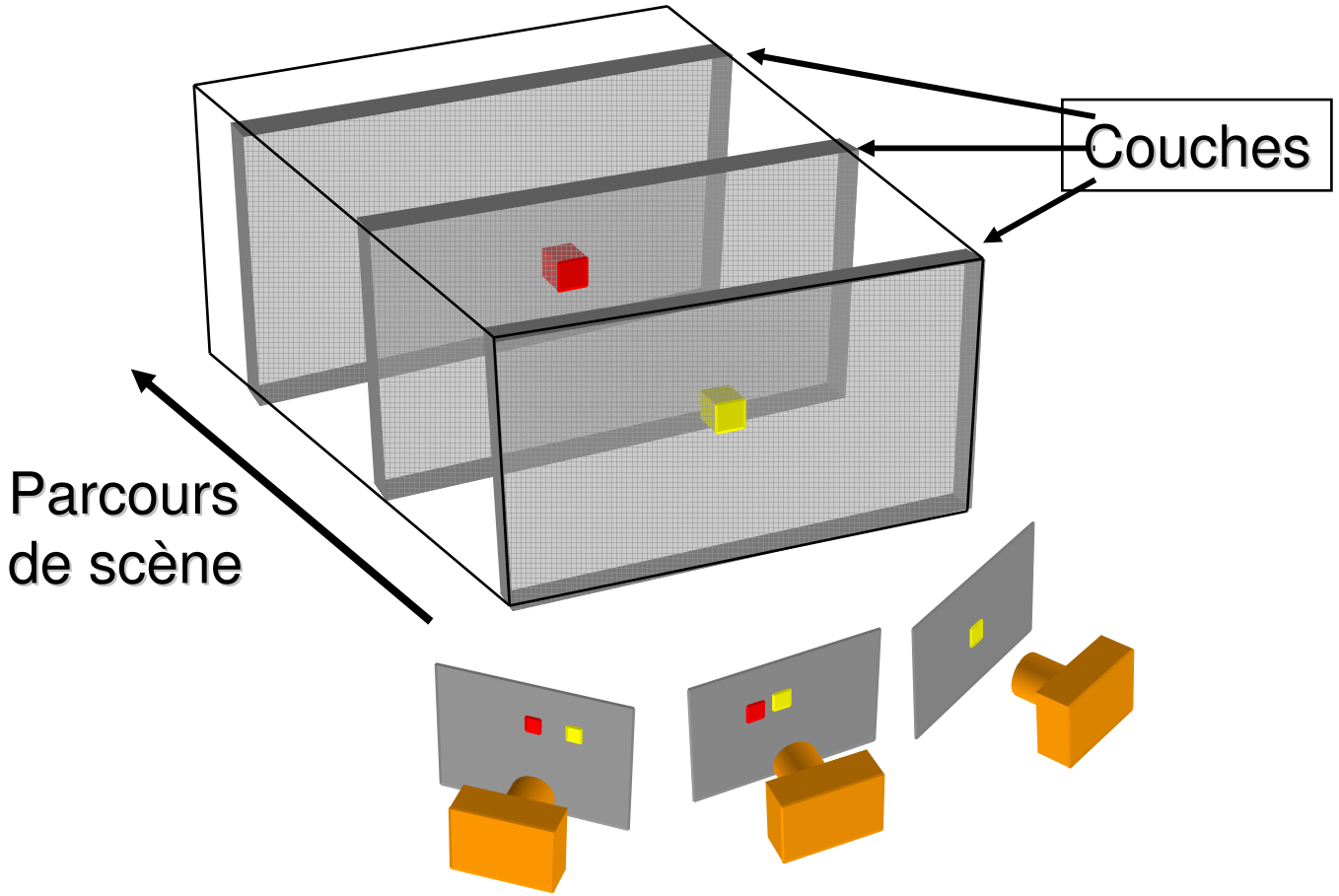
Visibilité directe



Scène inconnue

Visibilité inverse

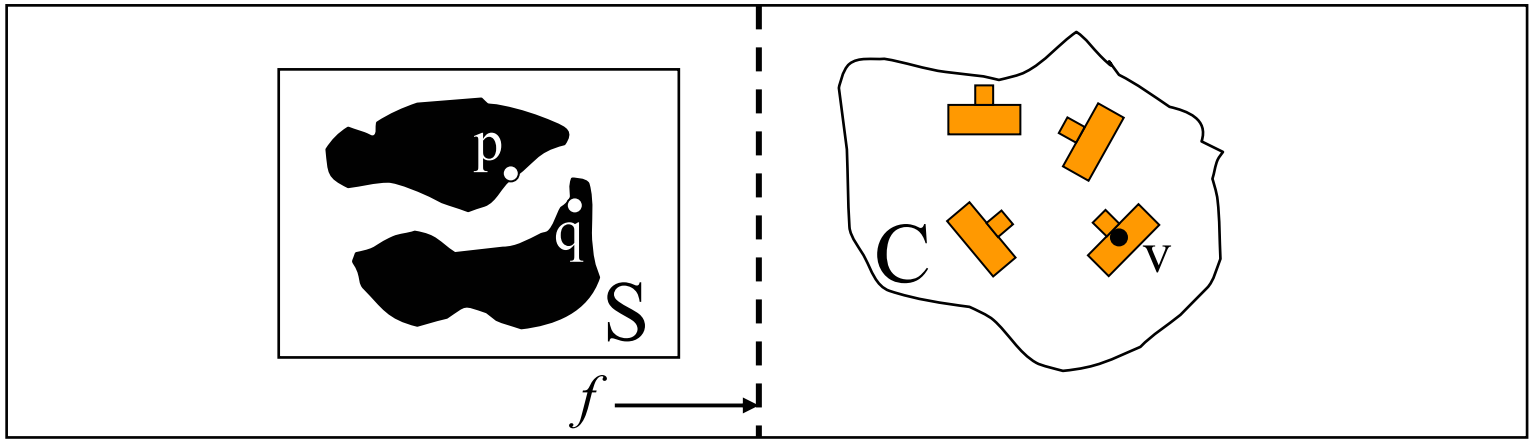
Ordre de profondeur: ceux qui cachent d'abord



Condition : l'ordre de profondeur doit être *indépendant du point de vue*

Ordre indépendant du point de vue

Une fonction f sur une scène S et un volume de caméra C



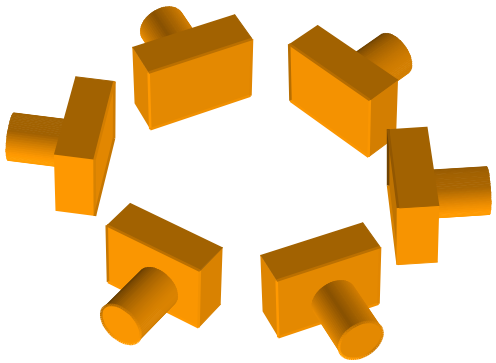
Tel que

p cache q de v *seulement si* f(p) < f(q)

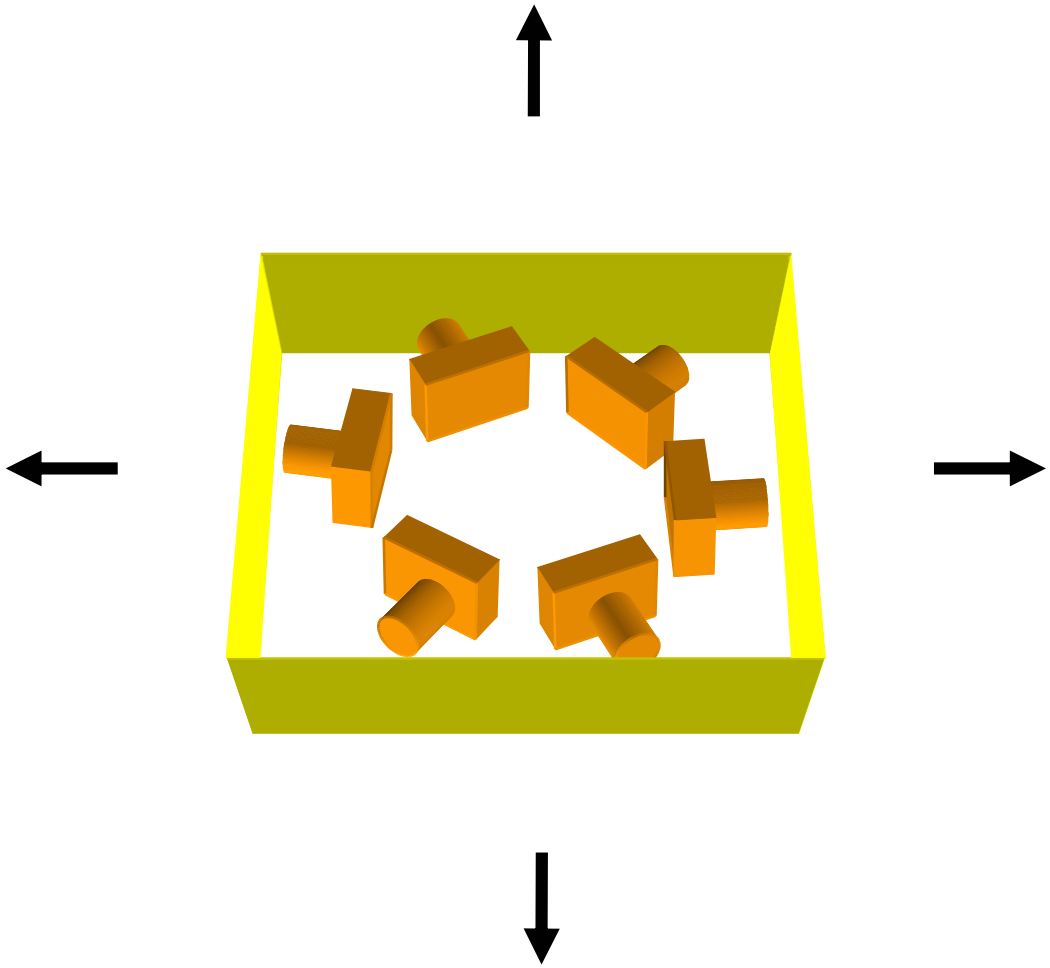
Par exemple : f = *distance au plan de séparation*

Ordre de profondeur panoramique

- Caméras orientées dans des directions différentes
- L'ordre de profondeur utilisant un plan ne fonctionne pas

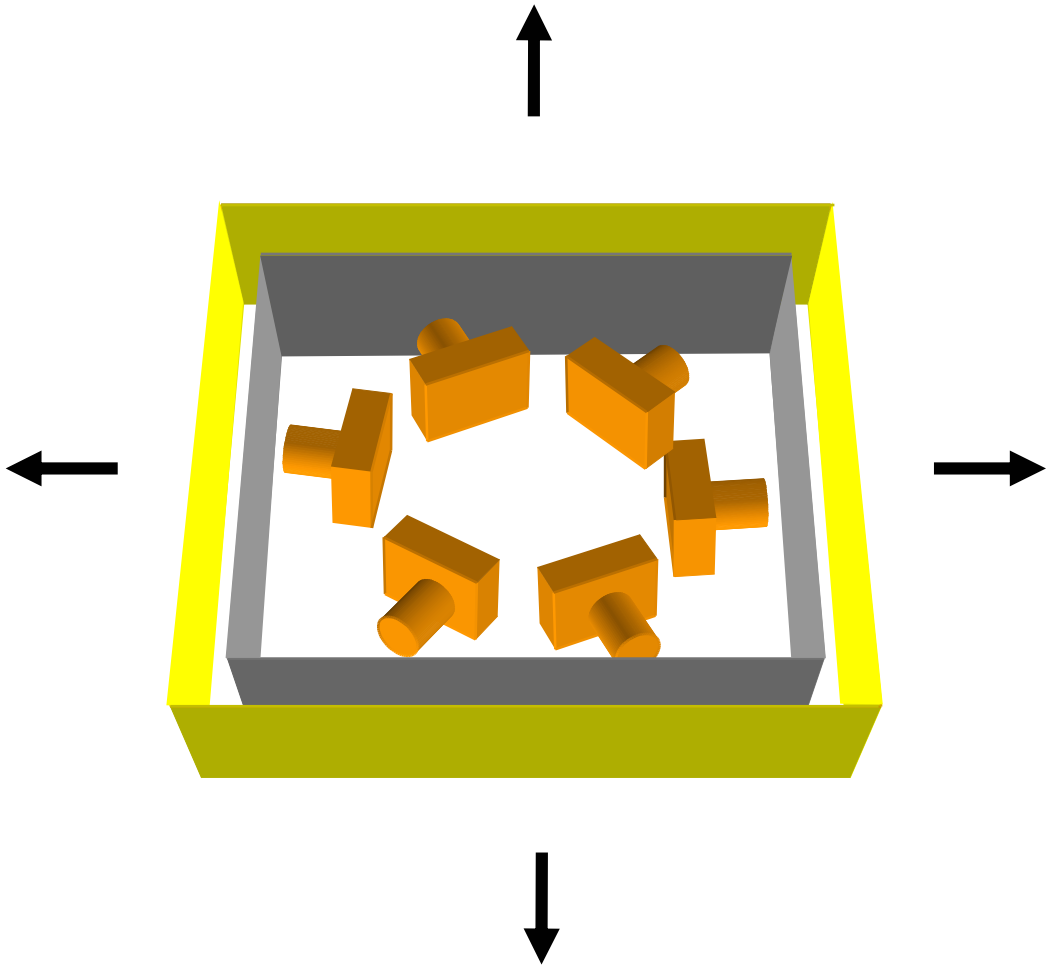


Ordre de profondeur panoramique



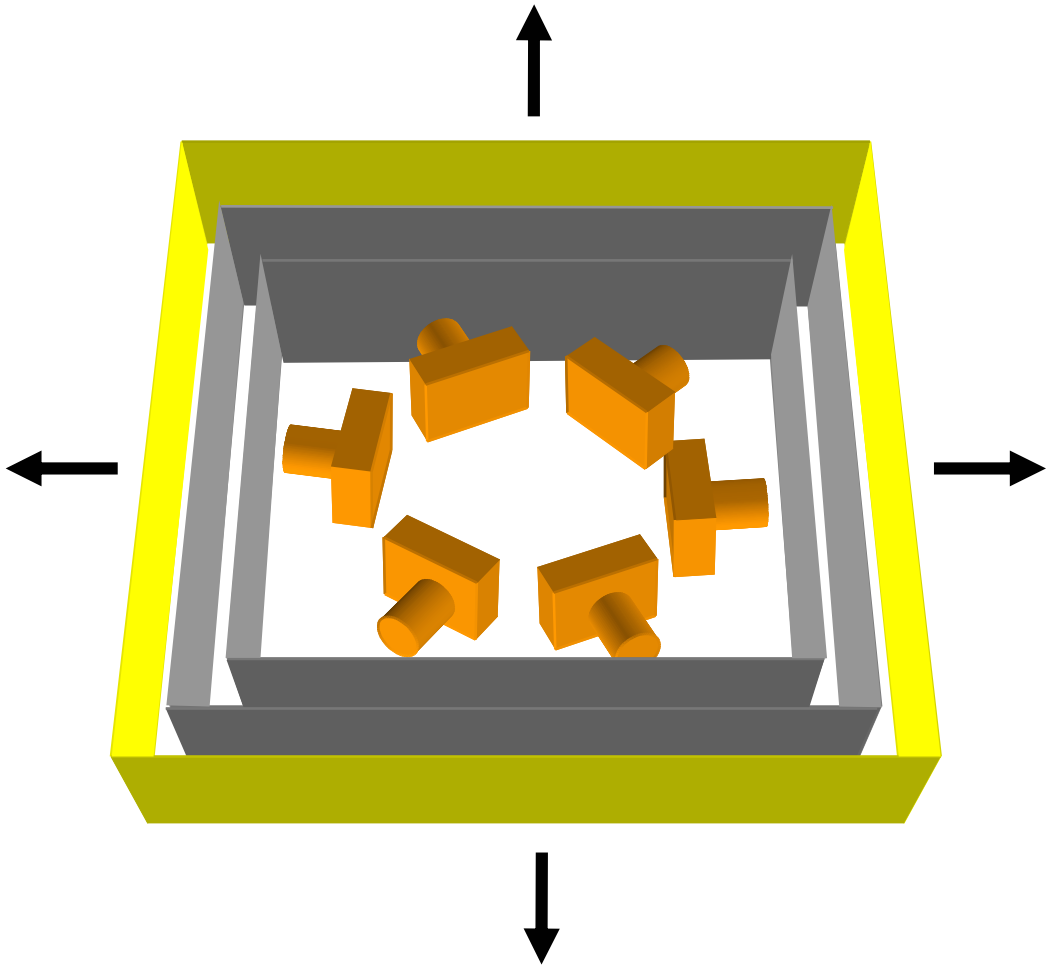
Les couches vont vers l'extérieur des caméras

Ordre de profondeur panoramique



Les couches vont vers l'extérieur des caméras

Ordre de profondeur panoramique

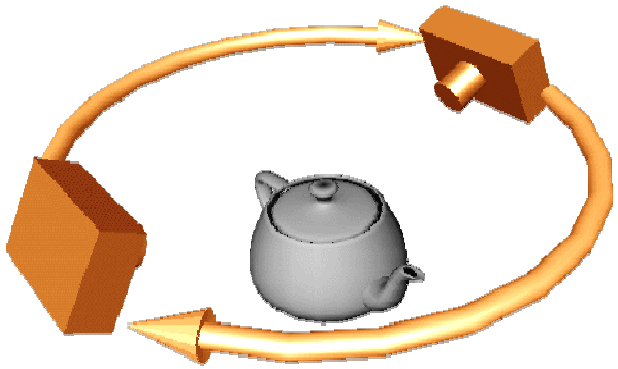


Les couches vont vers l'extérieur des caméras

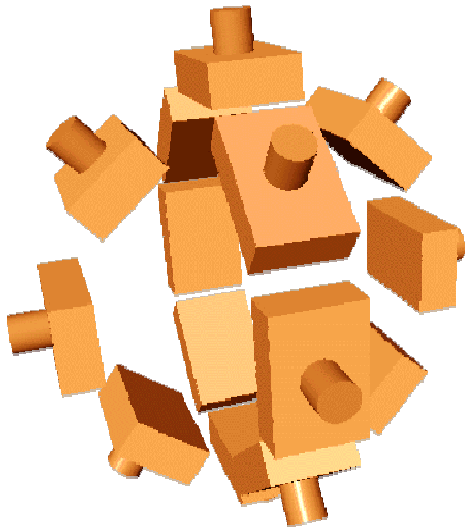
Configurations de caméras compatibles

Contraintes d'ordre de profondeur

- La scène doit être en dehors de l'enveloppe convexe des centres de caméras



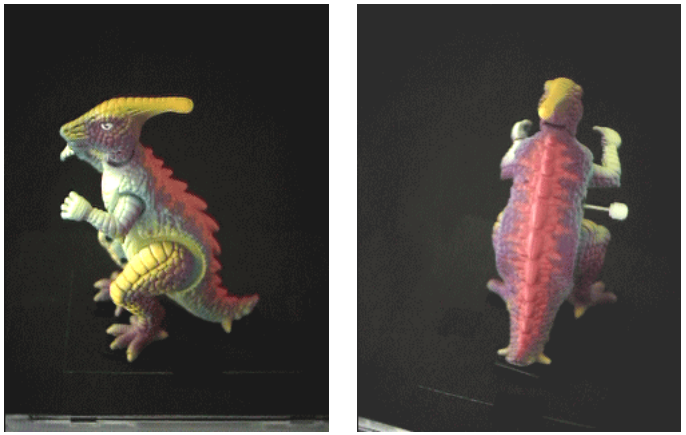
regard vers l'intérieur



regard vers l'extérieur

Acquisition d'images calibrées

table tournante calibrée



Images de dinosaure

Images de fleur

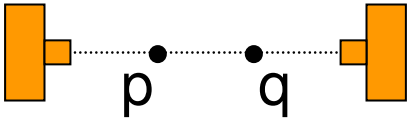
Résultats de coloriage de voxels

reconstruction de dinosaure
72 K voxels coloriés
7.6 M voxels testés
7 min. de calcul
sur SGI 250MHz

reconstruction de fleur
70 K voxels coloriés
7.6 M voxels testés
7 min. de calcul
sur SGI 250MHz

Limitations de l'ordre de profondeur

Un ordre de profondeur indépendant du point de vue peut ne pas exister



Besoin d'algorithme s'appliquant au cas général

- Positions de caméras non contraintes
- Géométrie / topologie de la scène non contrainte

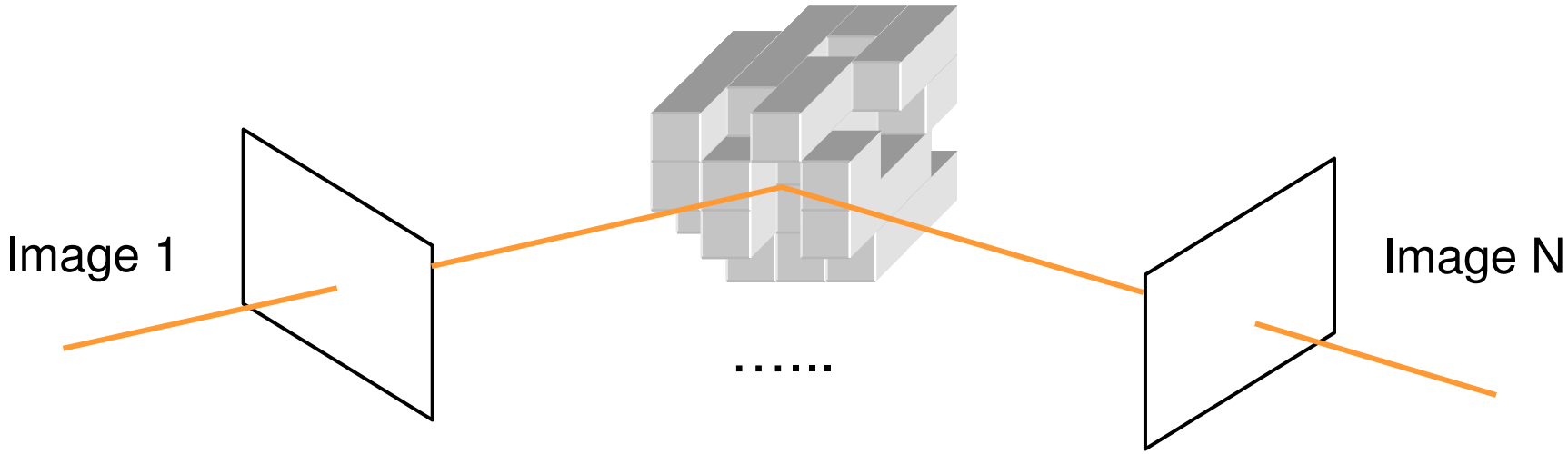
Solutions au coloriage de voxels

1. C=2 (silhouettes)
   - Intersection de volumes [Martin 81, Szeliski 93]

2. C non contraint, contraintes sur les points de vue
   - Algorithme de coloriage de voxels [Seitz & Dyer 97]

3. Cas général
   - Space carving (creusage de l'espace) [Kutulakos & Seitz 98]

Algorithme de space carving



Algorithme de space carving

- Initialiser à un volume V contenant la vraie scène
- Choisir un voxel sur la surface courante
- Projeter sur les images sur lesquelles il est visible
- Creuser si non photo-cohérent
- Répéter jusqu'à convergence

Convergence

Propriété de cohérence

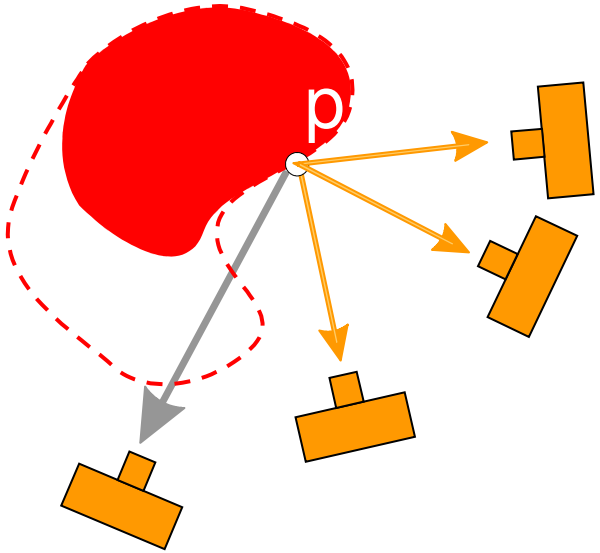
- La forme résultant est photo-cohérente

  > tous les points non cohérents ont été enlevés

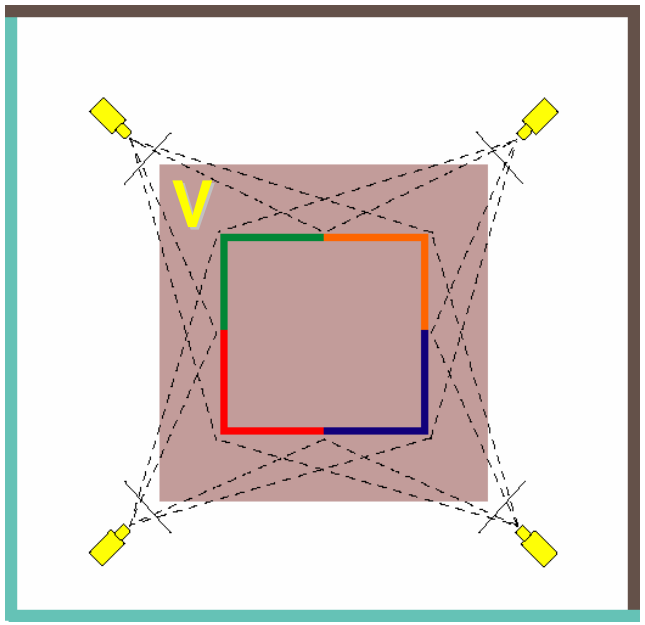
Propriété de convergence

- Le creusage converge vers un forme non-vide

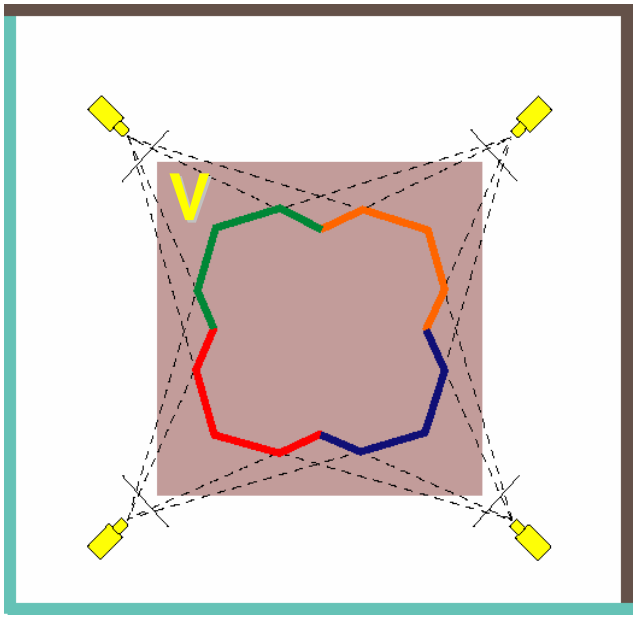
  > un point de la vraie scène n'est *jamais* enlevé



Qu'est-ce qu'on peut calculer ?



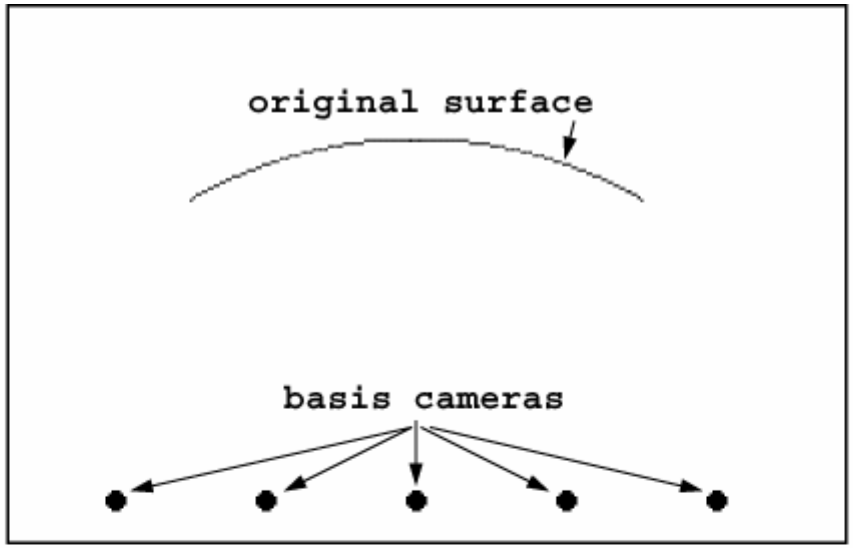
vraie scène



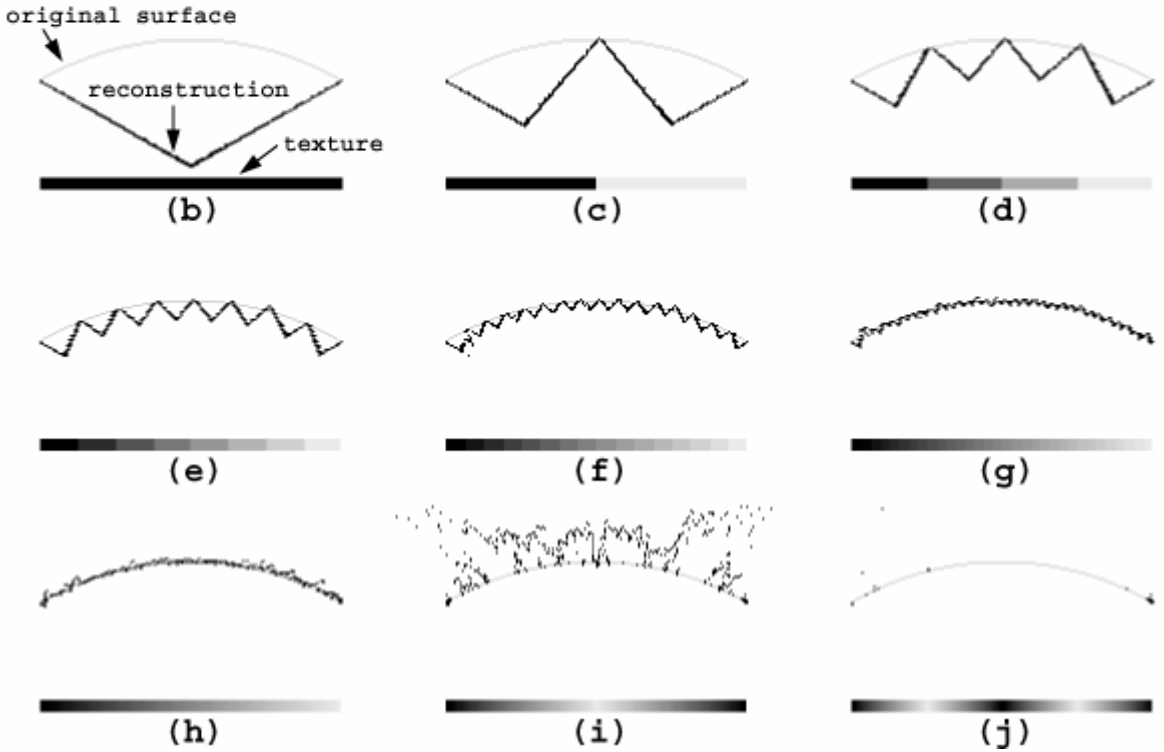
enveloppe photo

L'enveloppe photo *est l'union de toutes les scènes photo-cohérentes dans V*

- C'est une scène photo cohérente
- Plus petite borne possible sur la vraie scène


original surface
basis cameras

(a)


original surface
reconstruction
texture

(b) (c) (d)

(e) (f) (g)

(h) (i) (j)

Algorithme de *space carving*

L'algorithme de base est très compliqué

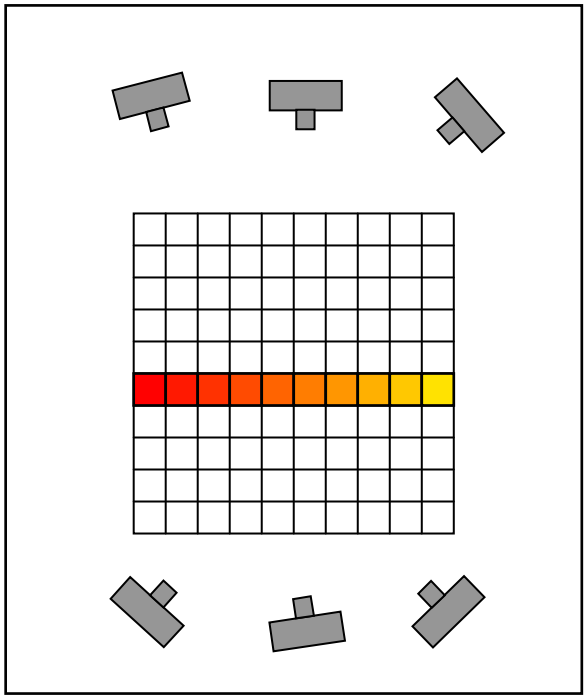
- Procédure de mise à jour complexe

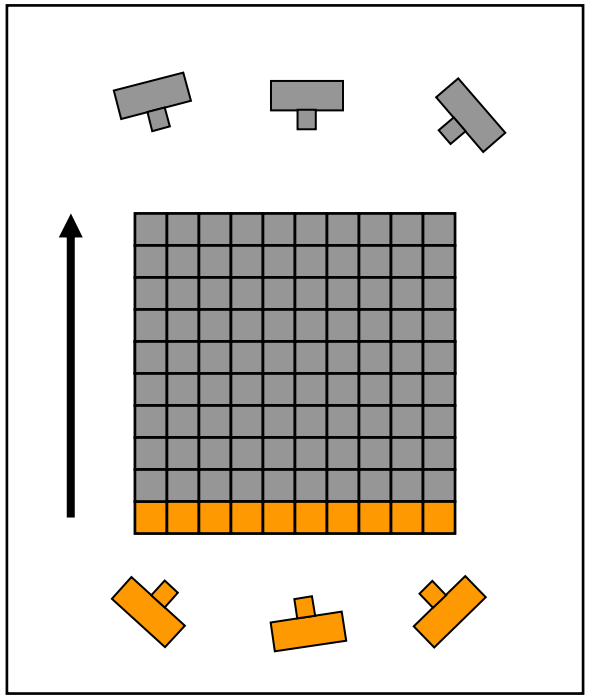
Alternative: Balayage de plan multi-passe

- Efficace, peut utiliser le texture-mapping hardware
- Converge rapidement en pratique
- Facile à implanter

Balayage de plan multi-passe

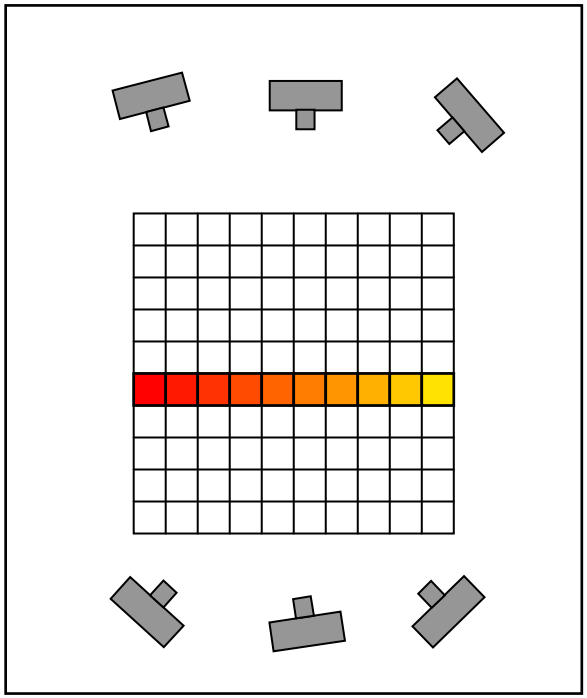
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

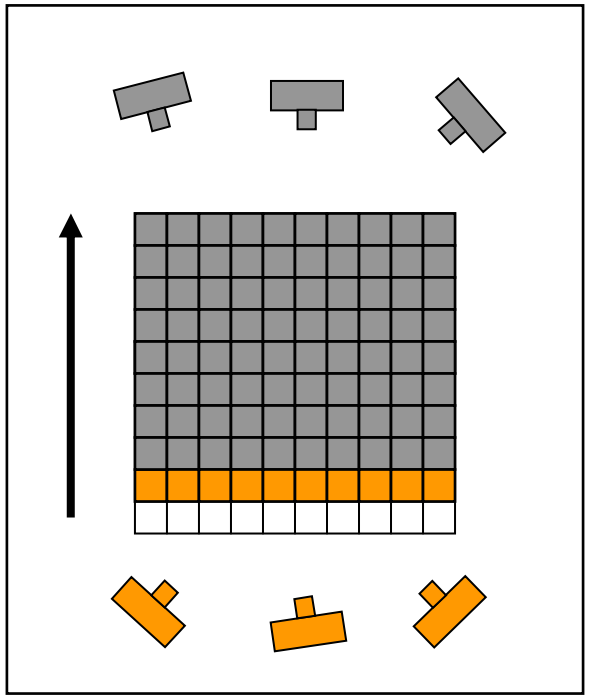




Balayage de plan multi-passe

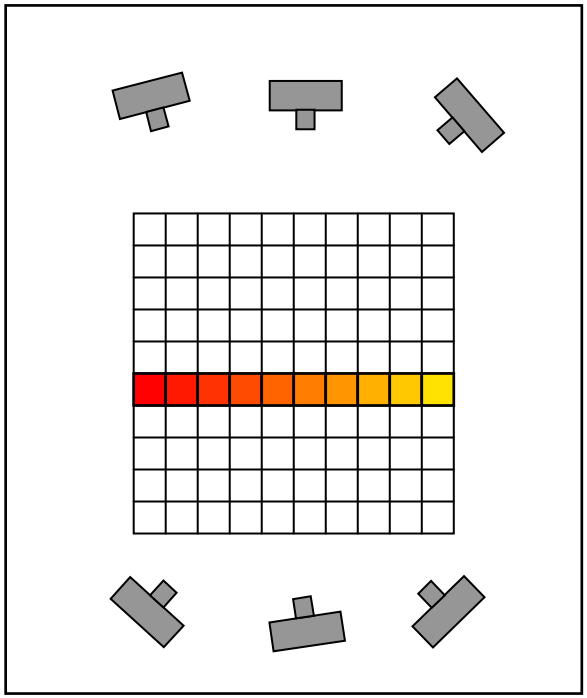
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

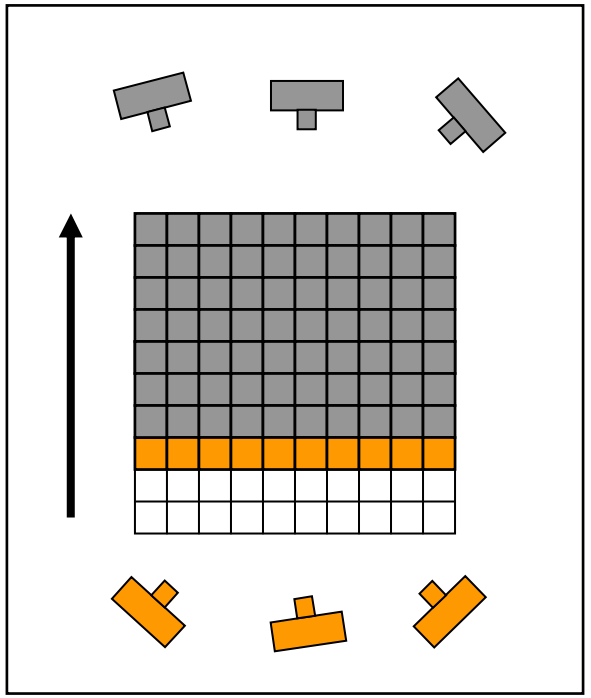




Balayage de plan multi-passe

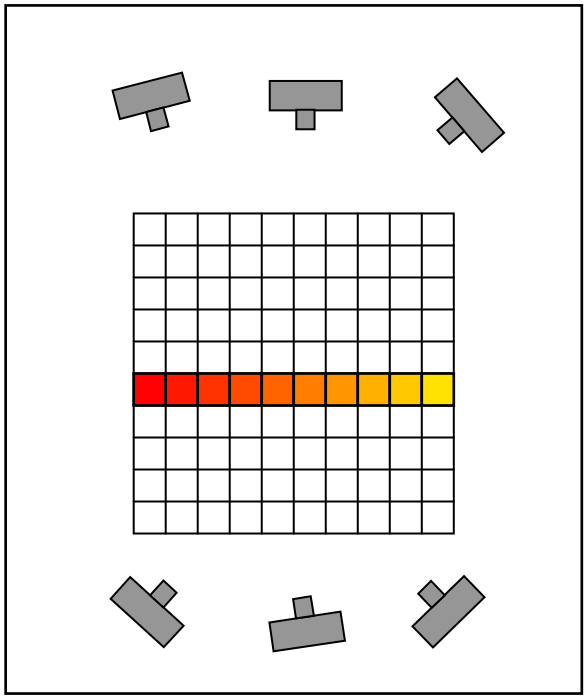
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

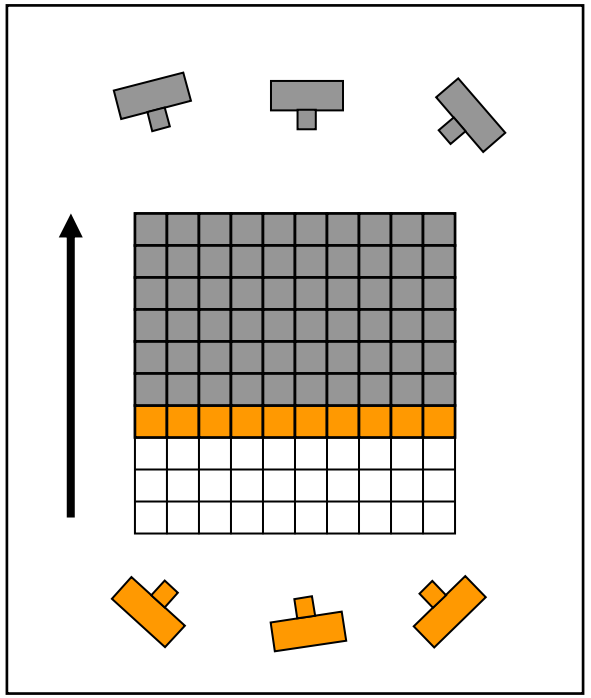




Balayage de plan multi-passe

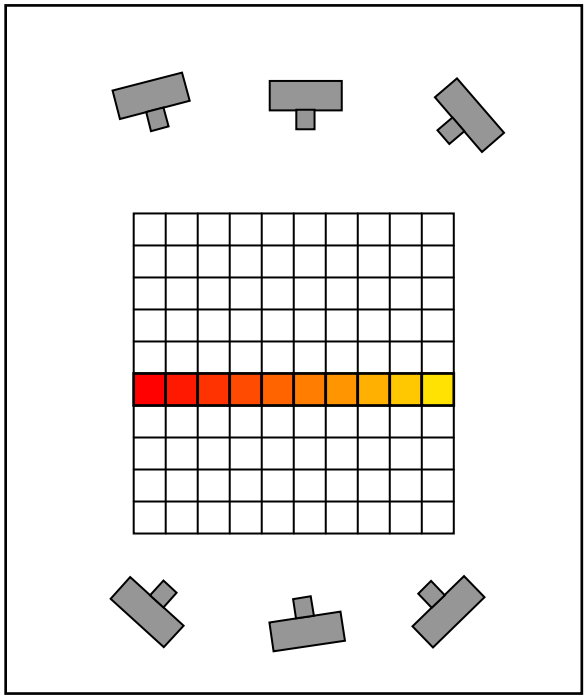
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

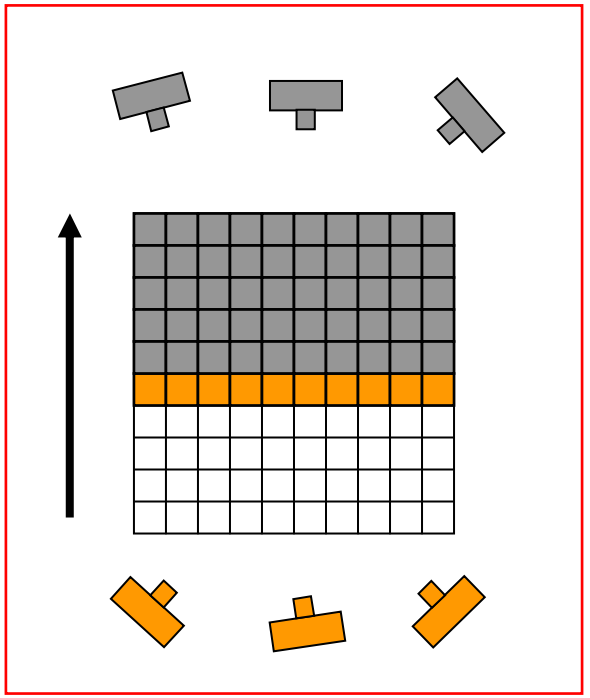




Balayage de plan multi-passe

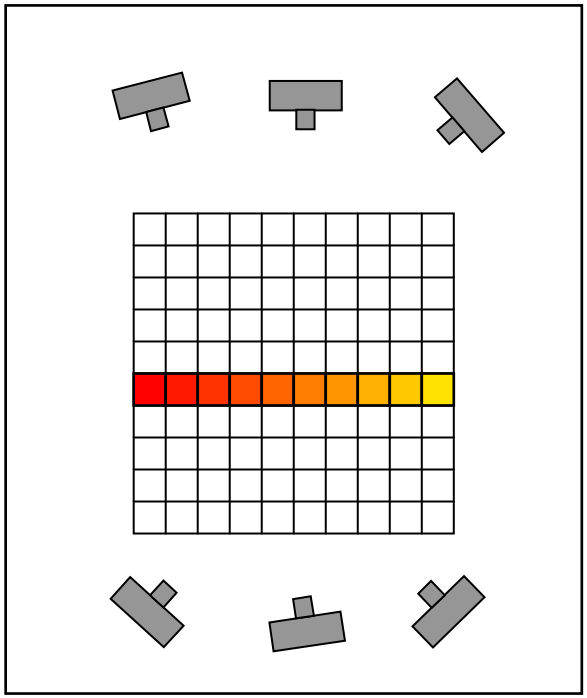
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

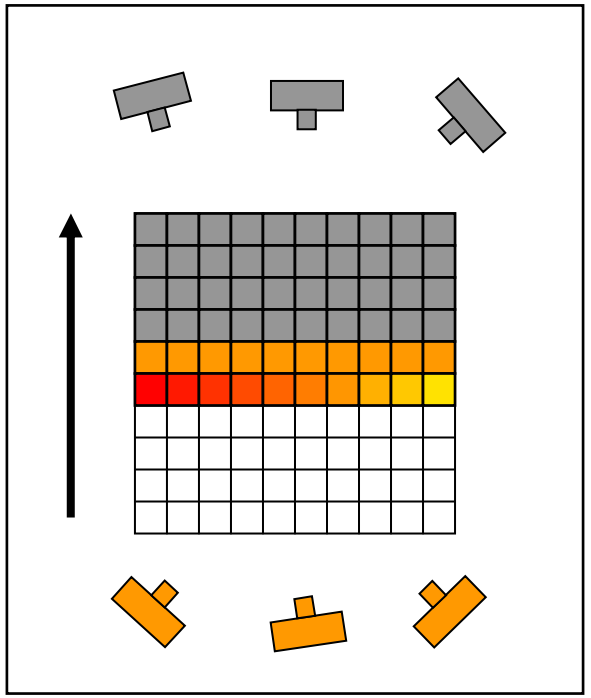




Balayage de plan multi-passe

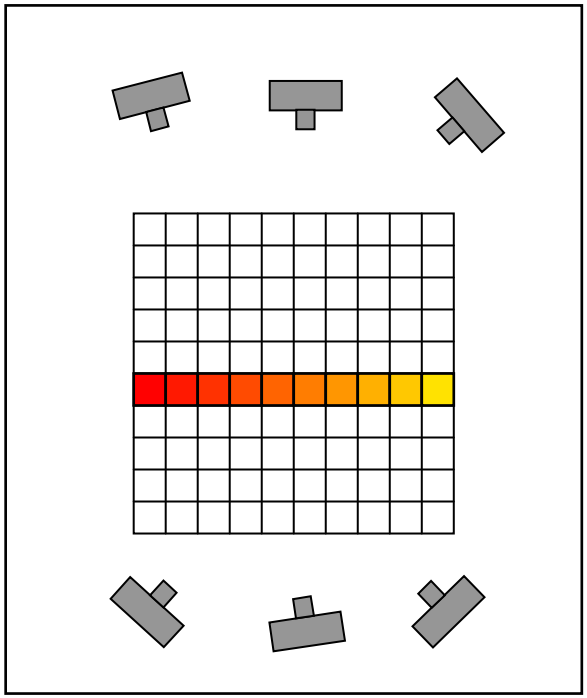
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

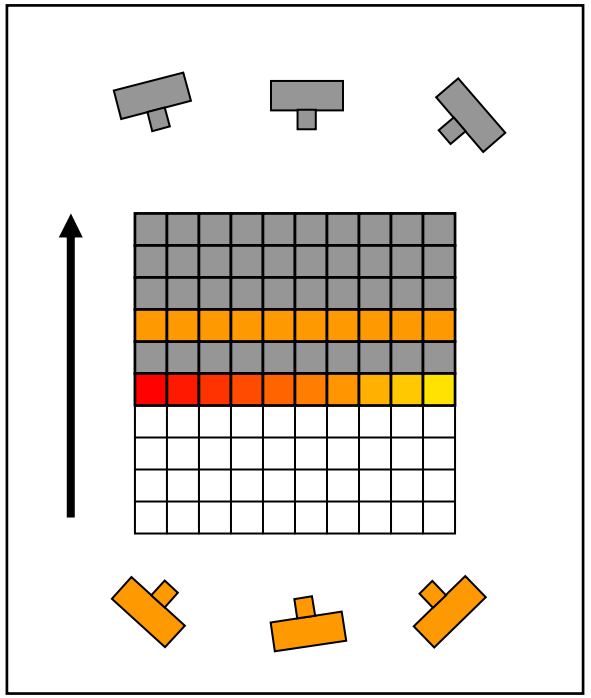




Balayage de plan multi-passe

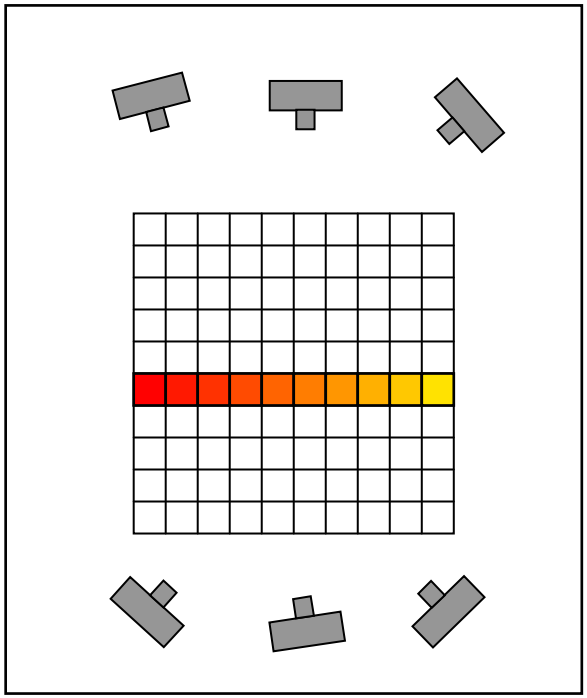
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

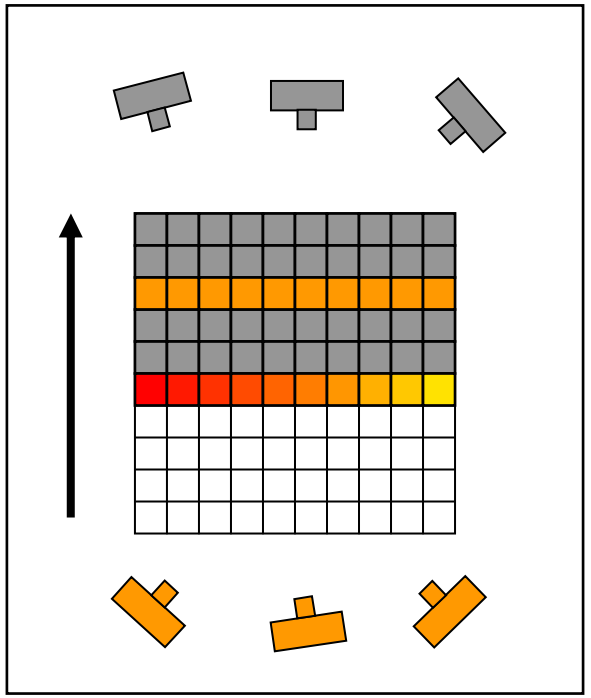




Balayage de plan multi-passe

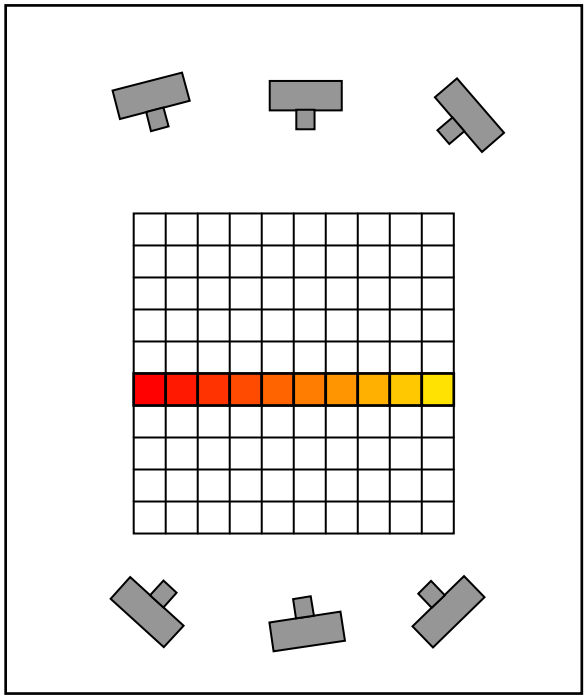
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

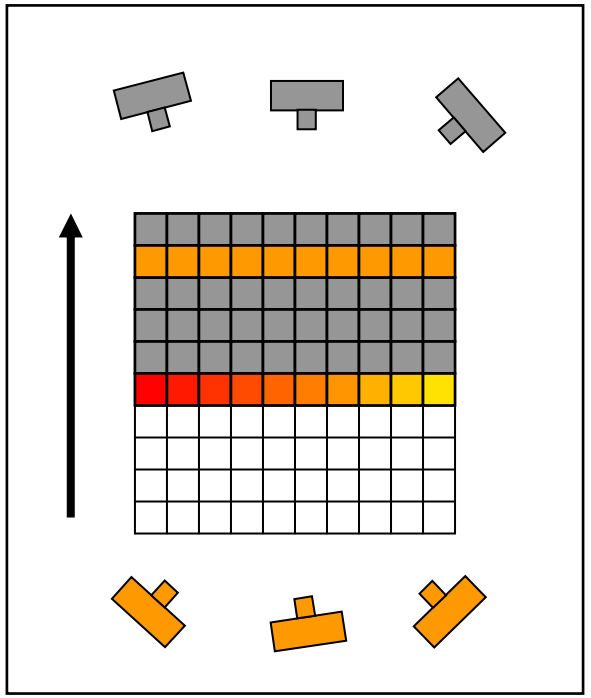




Balayage de plan multi-passe

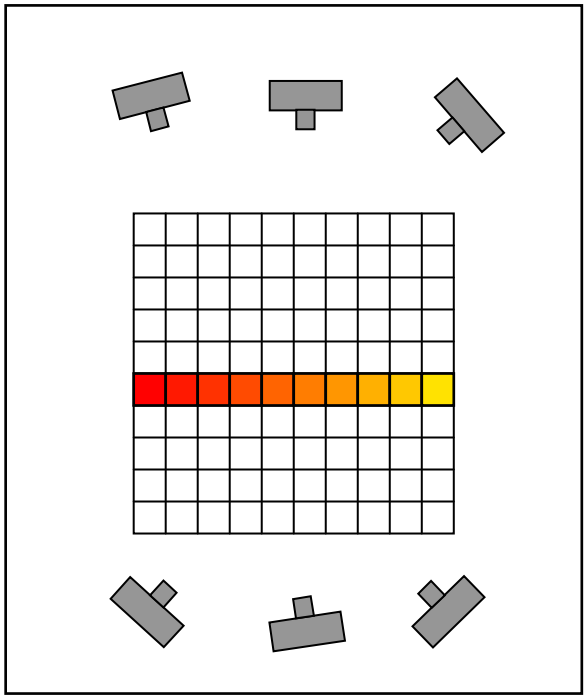
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

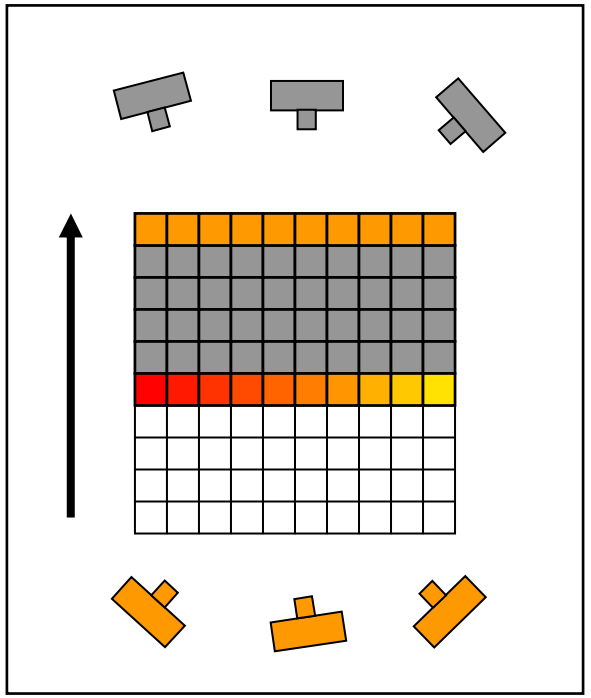




Balayage de plan multi-passe

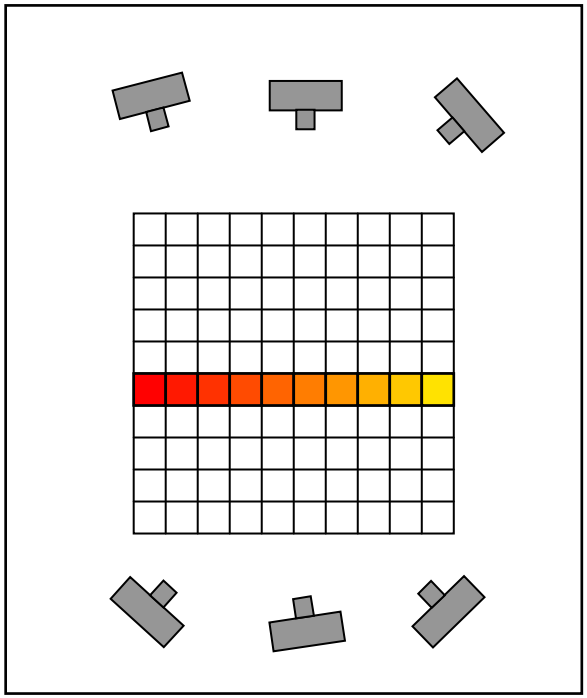
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

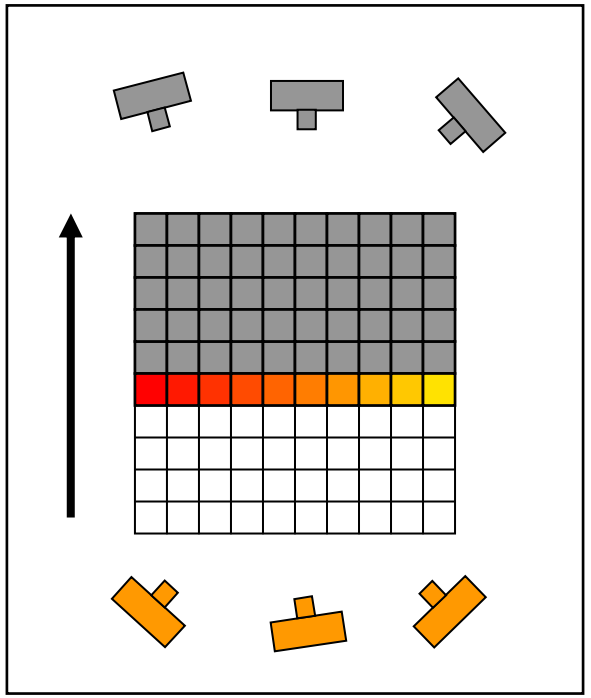




Balayage de plan multi-passe

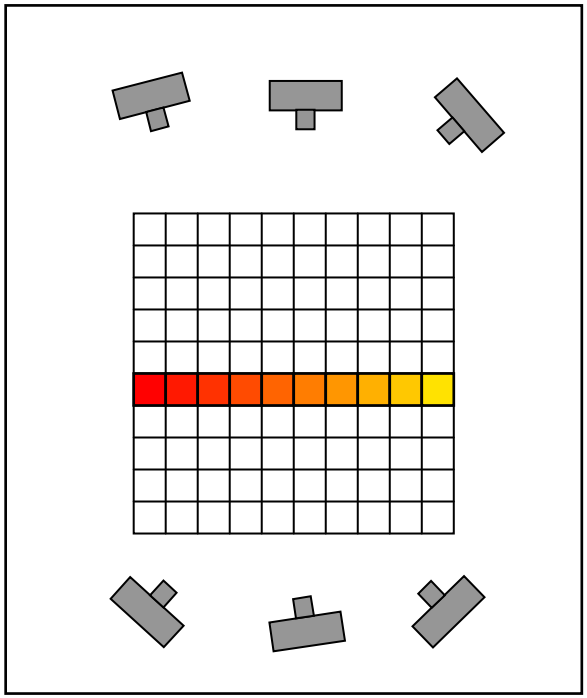
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

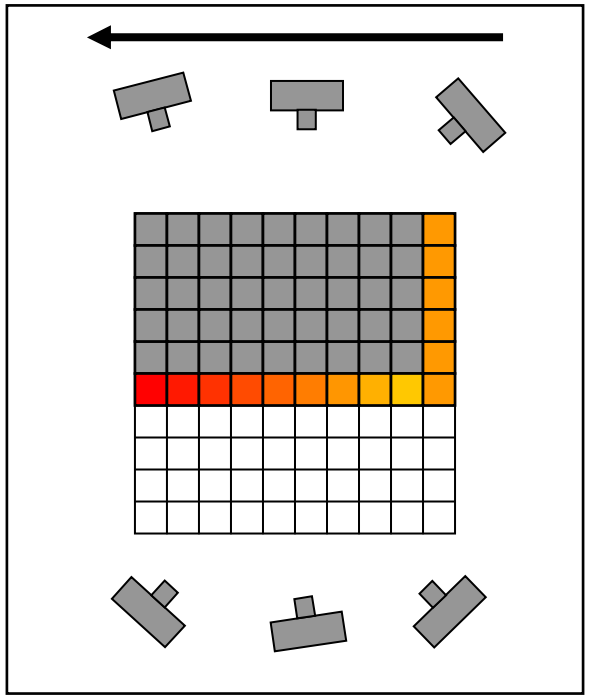




Balayage de plan multi-passe

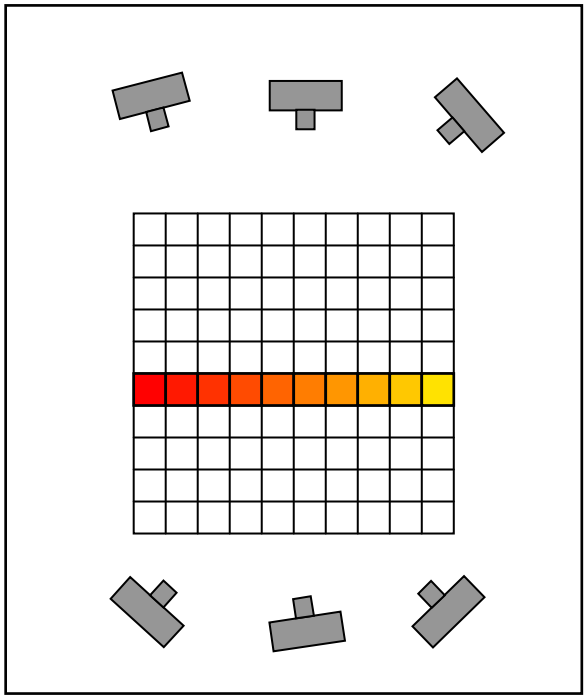
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

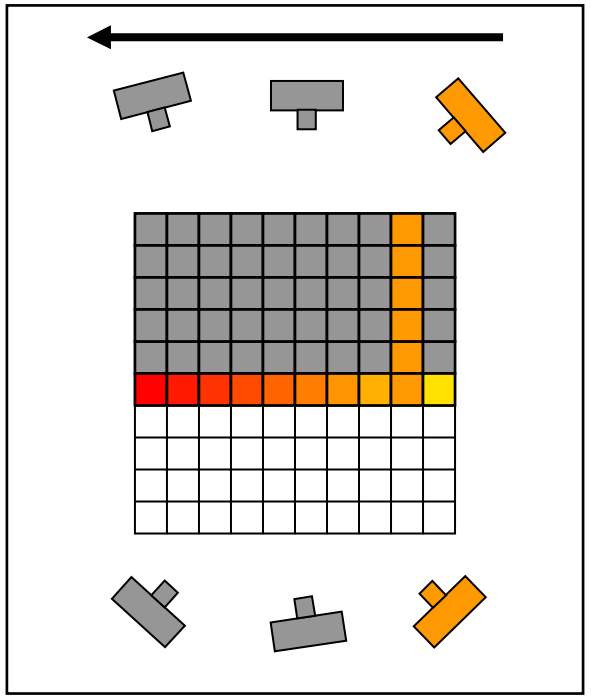




Balayage de plan multi-passe

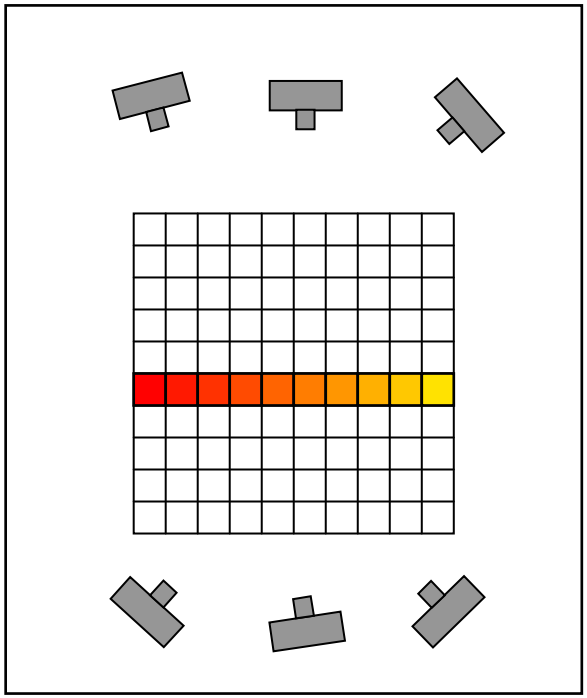
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

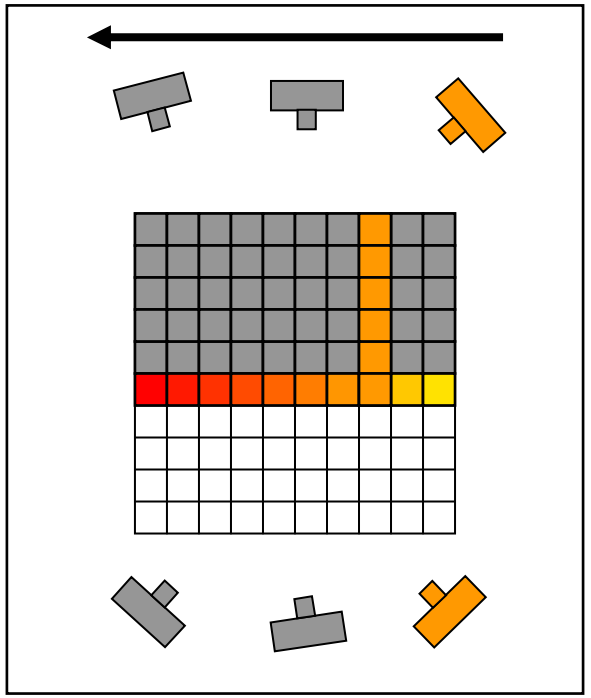




Balayage de plan multi-passe

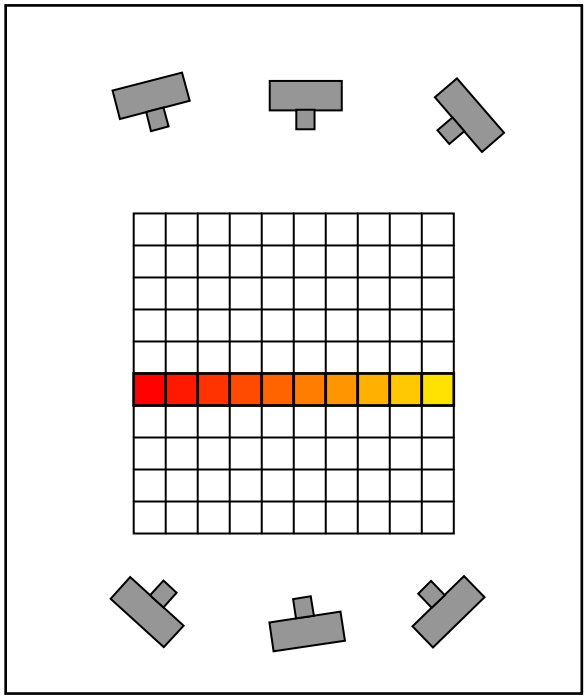
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

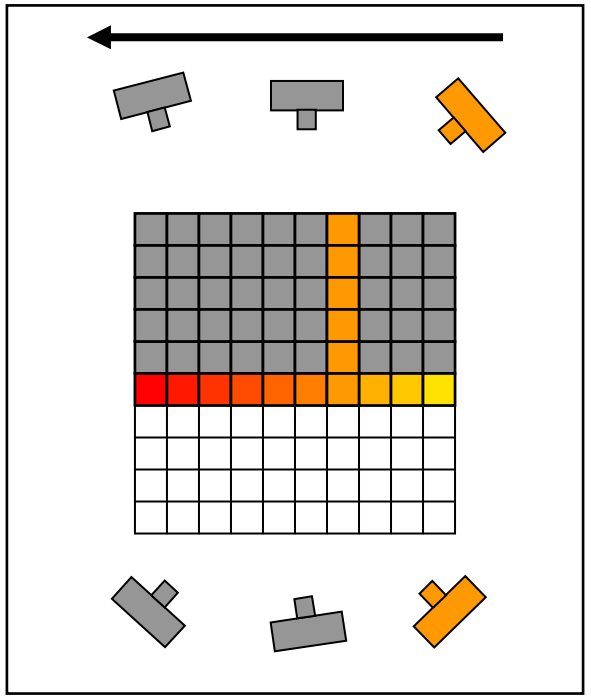




Balayage de plan multi-passe

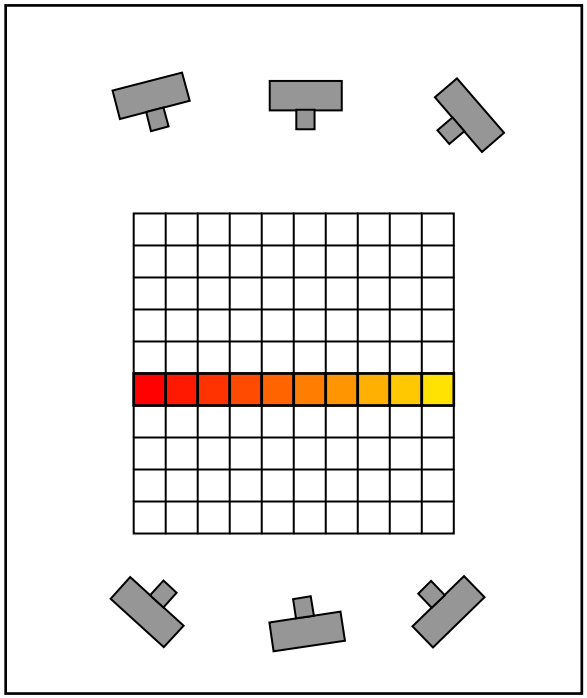
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

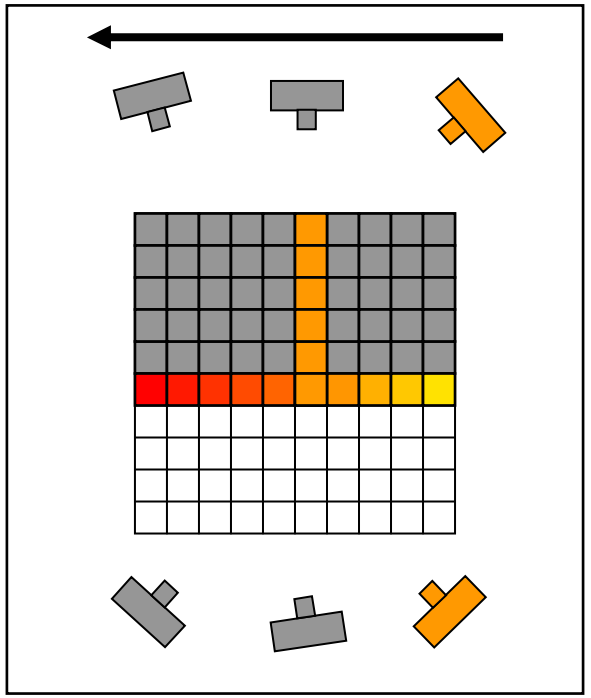




Balayage de plan multi-passe

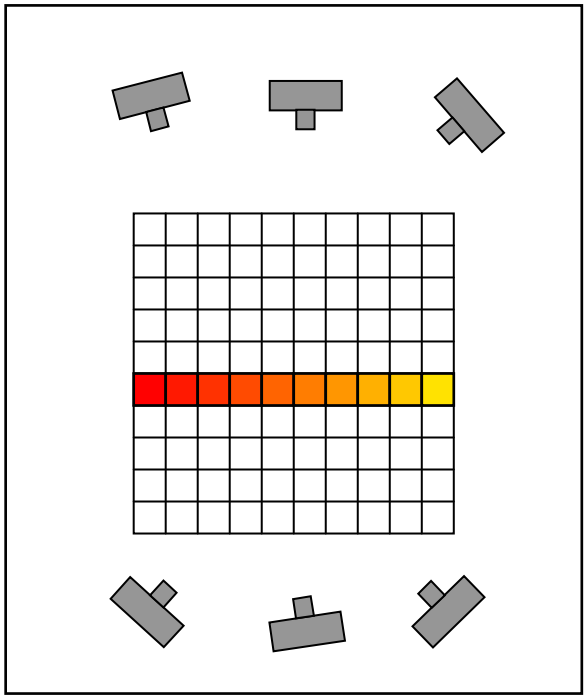
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

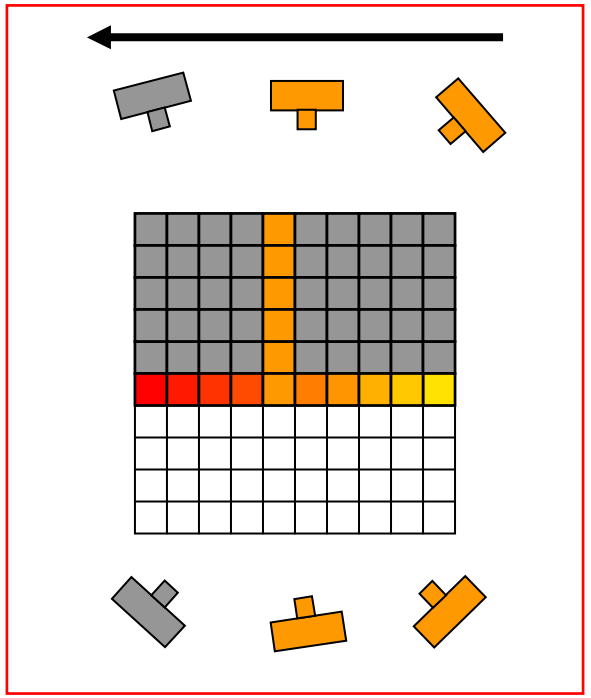




Balayage de plan multi-passe

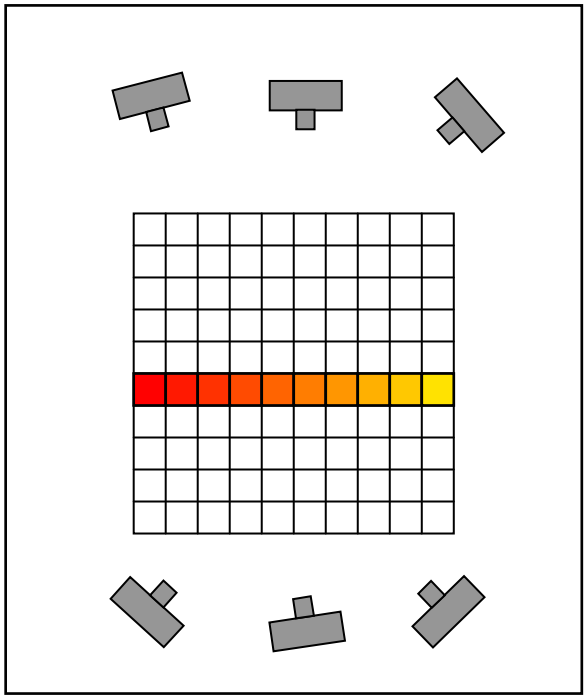
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

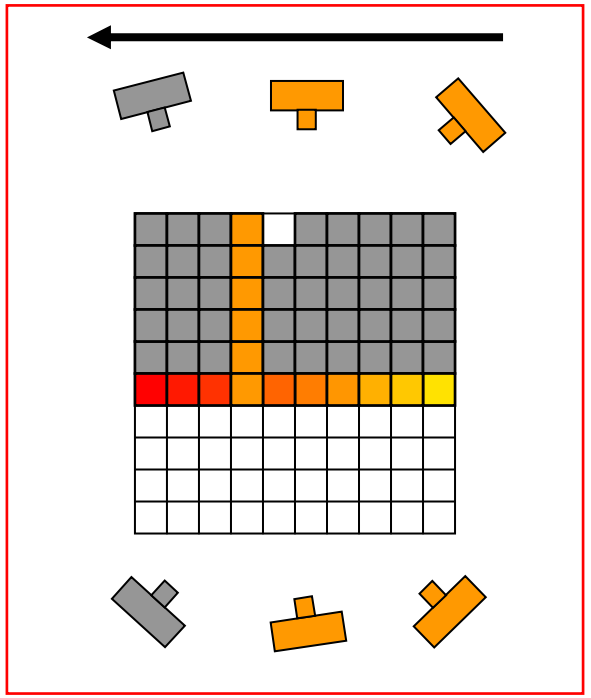




Balayage de plan multi-passe

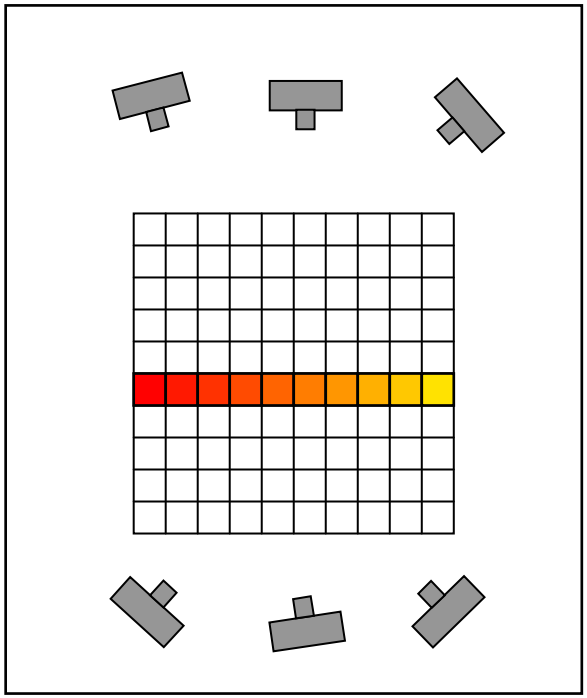
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

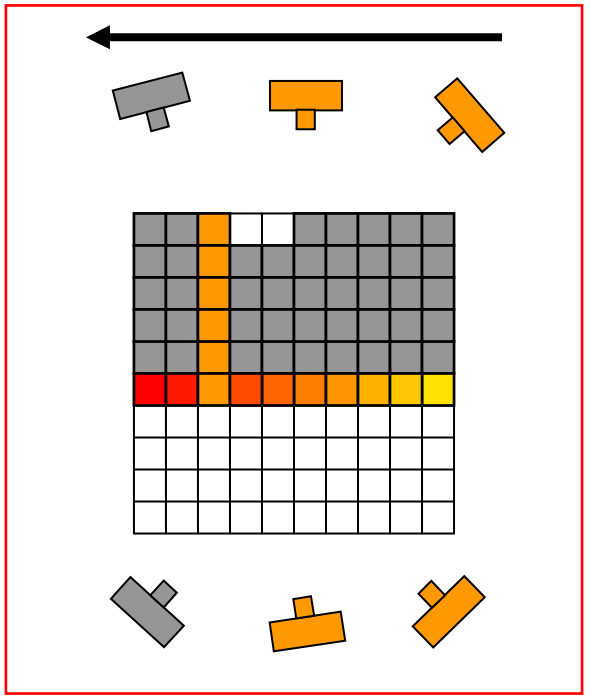




Balayage de plan multi-passe

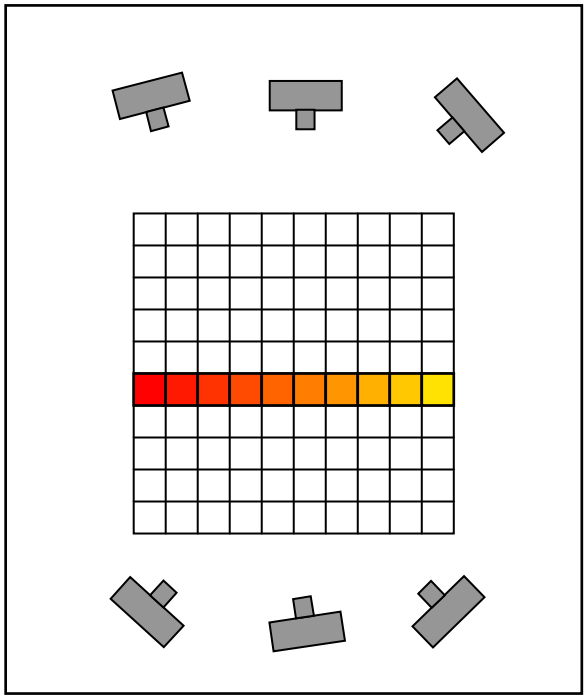
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

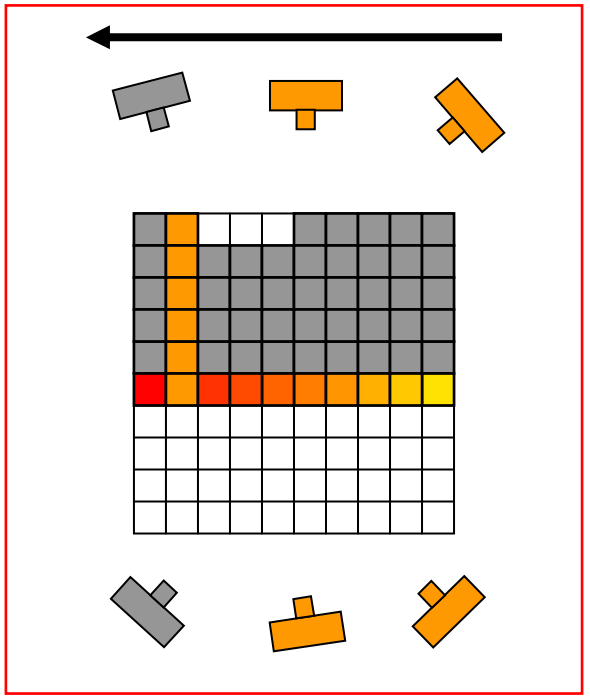




Balayage de plan multi-passe

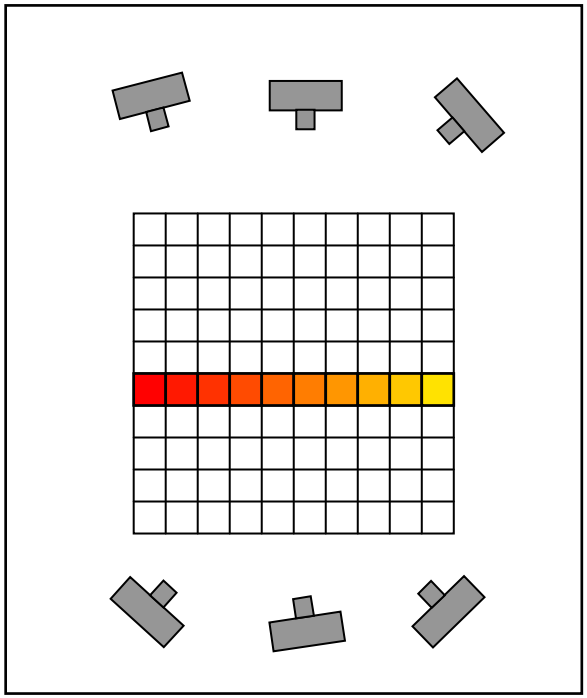
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

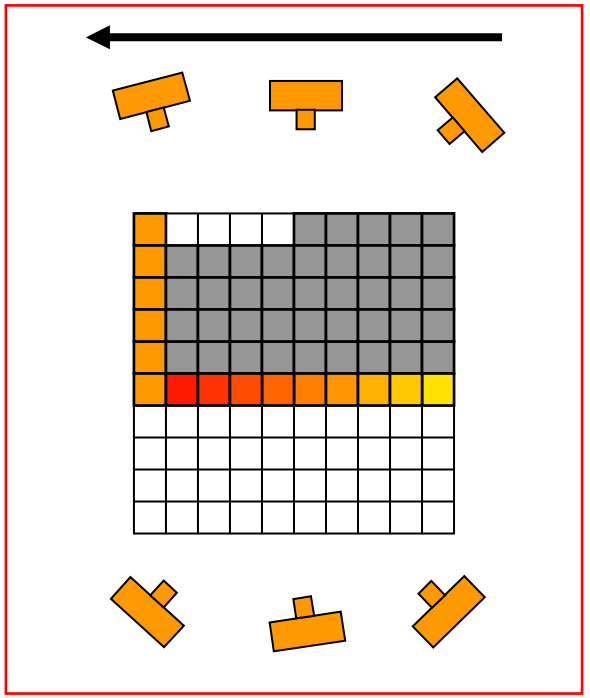




Balayage de plan multi-passe

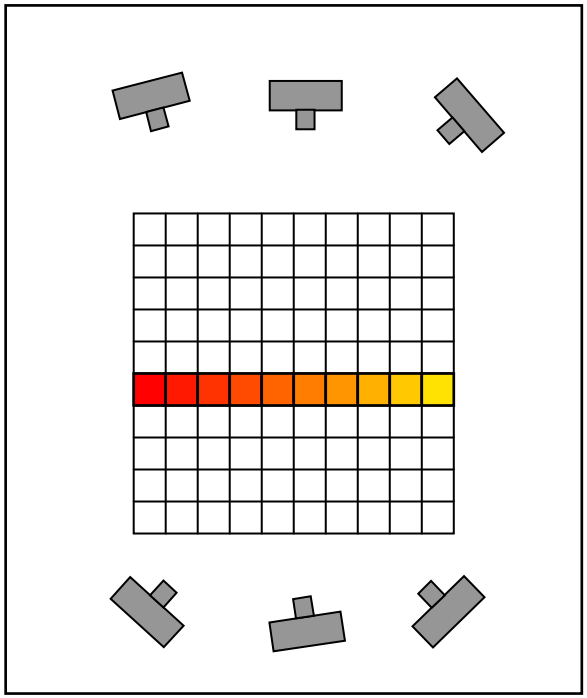
- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

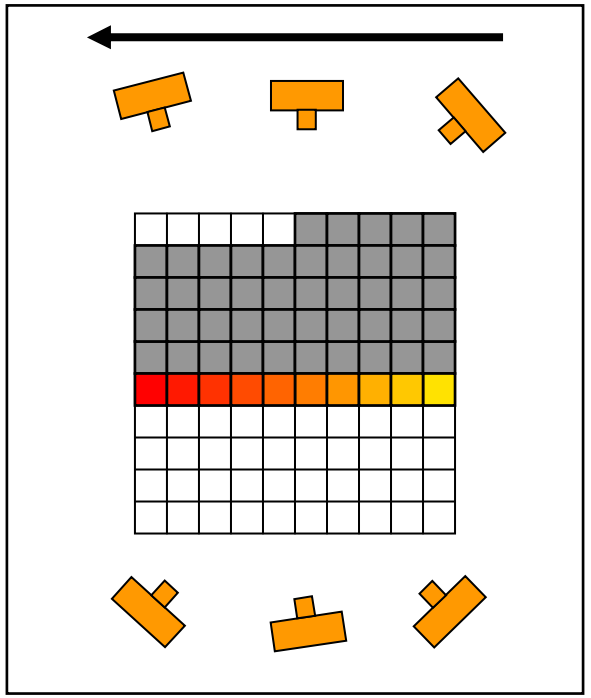




Balayage de plan multi-passe

- Balayer par un plan dans les 6 directions principales
- N'utiliser que les caméras d'un côté du plan
- Répéter jusqu'à convergence

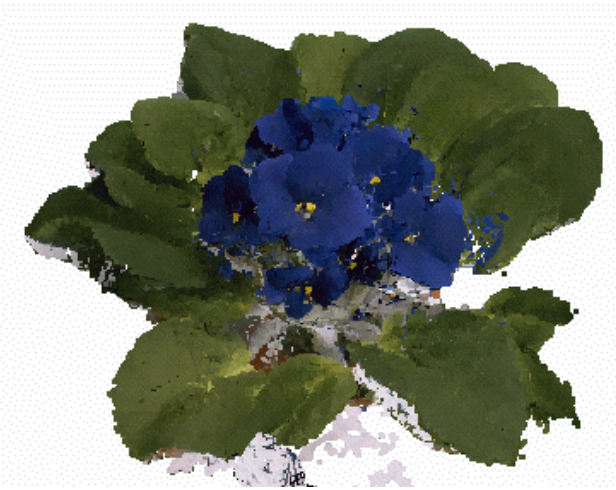




Résultats de *space carving* : violette africaine

Image d'entrée (1 sur 45)

Reconstruction



Reconstruction

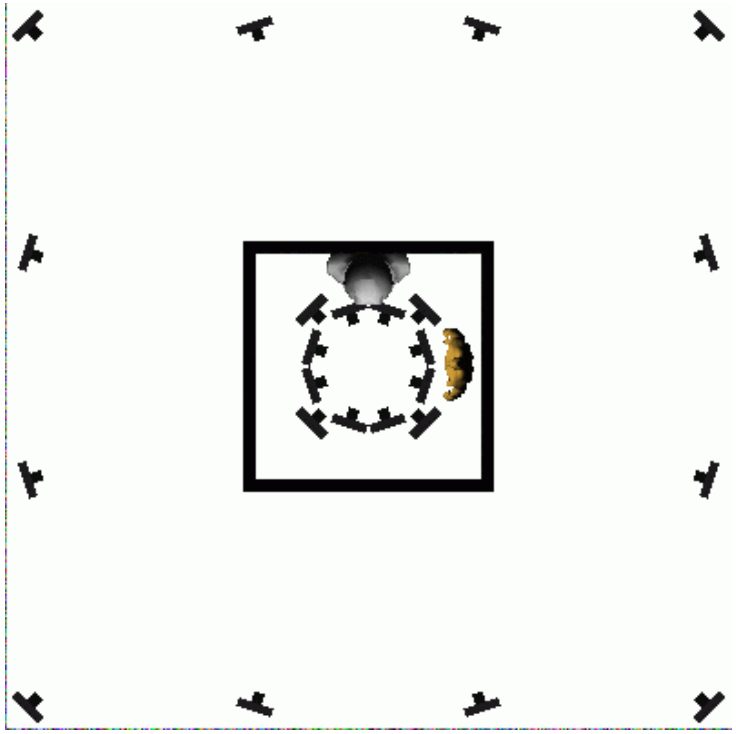
Reconstruction

Résultats de *space carving* : main

Image d'entrée
(1 sur 100)

Vues de la reconstruction

Ballade dans la maison



En entrée : 24 vues générées de l'intérieur *et* de l'extérieur

Résultats de *space carving* : maison

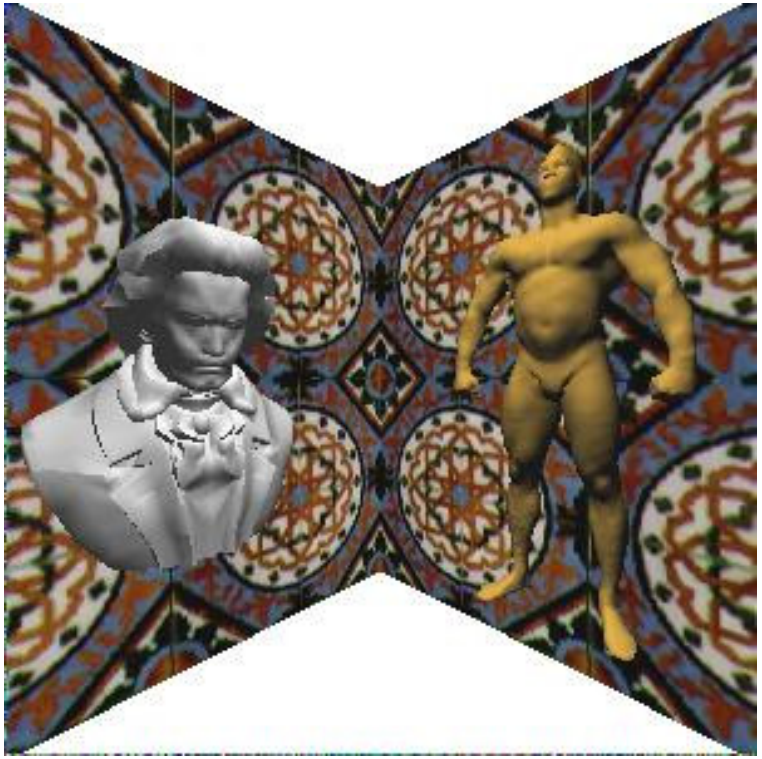


Image d'entrée
(vraie scène)

Reconstruction
370,000 voxels

Résultats de *space carving* : maison

Image d'entrée
(vraie scène)

Reconstruction
370,000 voxels

Résultats de *space carving* : maison

autre vue (vraie scène)

Reconstruction

autre vue (vraie scène)

Reconstruction

Reconstruction (en utilisant l'autre vue)

Autres caractéristiques

Reconstruction grossière puis fine

- Representation de la scène comme un octree
- Reconstruction du modèle basse résolution d'abord, puis affiner

Accélération matérielle

- Utilisation du texture-mapping pour la projection des voxels
- Voxels traités plan par plan

Limitations

- Acquisition d'images calibrées
- Restriction à des modèles de radiance simples
- Biais vers les reconstructions maximales (plus grosses)
- Pas de transparence

Autres approches

Ensembles de niveau *[Faugeras & Keriven 1998]*

- Évolution d'une fonction implicite par résolution d'EDP

Transparence *[Szeliski & Golland 1998]*

- Calcul des voxels avec une composante alpha

Flot maximal/Coupure minimale *[Roy & Cox 1998]*

- Formulation utilisant la théorie des graphes

Stéréo basée sur un maillage *[Fua & Leclerc 95]*

- Basée sur un maillage mais même notion de cohérence

Virtualized Reality *[Narayan, Rander, Kanade 1998]*

- Stéréo sur de multiples triplets, fusion des résultats

Stéréo par ensembles de niveau

Stéréo par minimisation d'énergie

- Première idée : meilleur surface S(u,v) qui correspond aux images

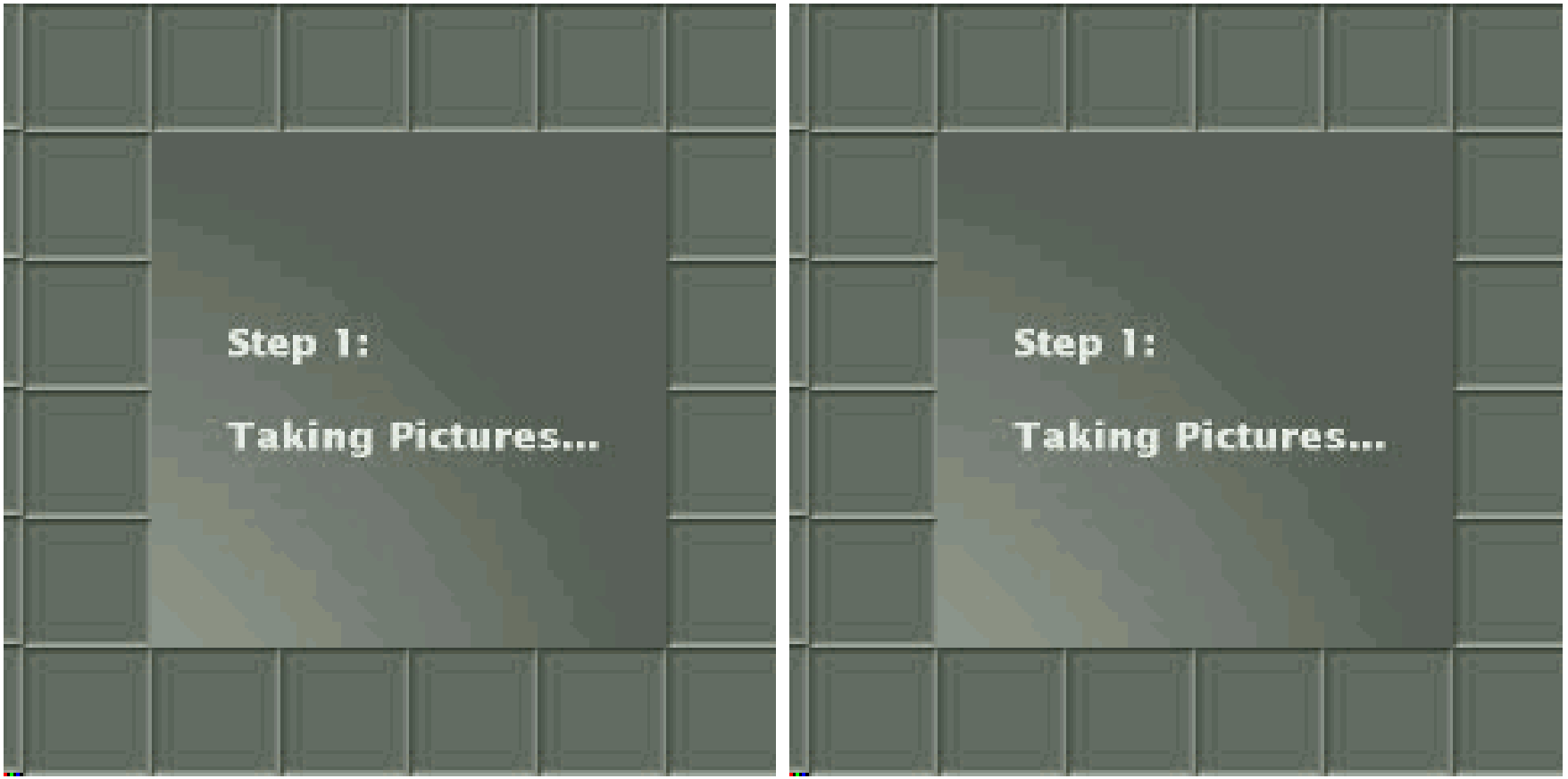
$$E = \int_u \int_v photoconsistency(S(u,v))dudv$$

- C'est un problème de minimisation variationnel
  > déformations infinitésimale d'une surface
  > deformation par équations d'Euler-Lagrange

Problème : et si l'objet n'a pas une surface unique

- **Utilisation de la formulation par ensembles de niveau**
  > la surface de la scène est $S = \{ (x,y,z) \mid f(x,y,z) = 0\}$
  > f est une fonction "distance signée" à la surface
  > on fait évoluer f par une EDP, plutôt que S
  > permet de gérer des topologies complexes

Stéréo par ensembles de niveau : résultats



Bibliographie

Intersection de volumes

- Martin & Aggarwal, “Volumetric description of objects from multiple views”, Trans. Pattern Analysis and Machine Intelligence, 5(2), 1991, pp. 150-158.
- Szeliski, “Rapid Octree Construction from Image Sequences”, Computer Vision, Graphics, and Image Processing: Image Understanding, 58(1), 1993, pp. 23-32.

Coloriage de voxels et space carving

- Seitz & Dyer, “Photorealistic Scene Reconstruction by Voxel Coloring”, Proc. Computer Vision and Pattern Recognition (CVPR), 1997, pp. 1067-1073.
- Seitz & Kutulakos, “Plenoptic Image Editing”, Proc. Int. Conf. on Computer Vision (ICCV), 1998, pp. 17-24.
- Kutulakos & Seitz, “A Theory of Shape by Space Carving”, Proc. ICCV, 1998, pp. 307-314.

Bibliographie

Autres méthodes

- Bolles, Baker, and Marimont, “Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion”, International Journal of Computer Vision, vol 1, no 1, 1987, pp. 7-55.
- Faugeras & Keriven, “Variational principles, surface evolution, PDE's, level set methods and the stereo problem", IEEE Trans. on Image Processing, 7(3), 1998, pp. 336-344.
- Szeliski & Golland, “Stereo Matching with Transparency and Matting”, Proc. Int. Conf. on Computer Vision (ICCV), 1998, 517-524.
- Roy & Cox, “A Maximum-Flow Formulation of the N-camera Stereo Correspondence Problem”, Proc. ICCV, 1998, pp. 492-499.
- Fua & Leclerc, “Object-centered surface reconstruction: Combining multi-image stereo and shading", International Journal of Computer Vision, 16, 1995, pp. 35-56.
- Narayanan, Rander, & Kanade, “Constructing Virtual Worlds Using Dense Stereo”, Proc. ICCV, 1998, pp. 3-10.