

# Master IGMMV

## Synthèse d'images et de sons

George Drettakis

Nicolas Tsingos



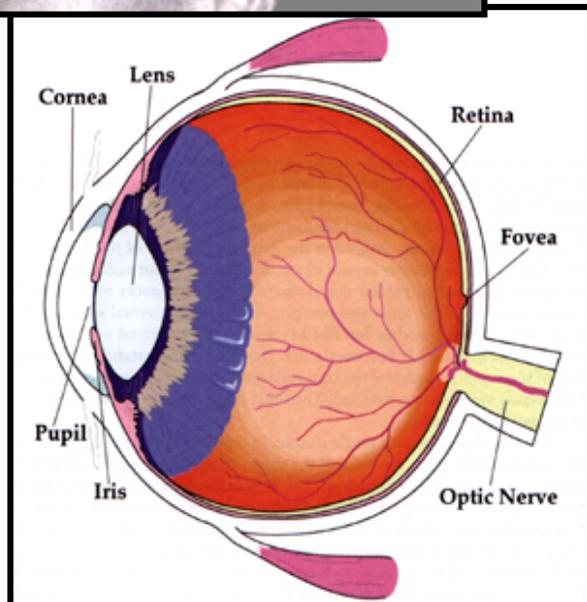
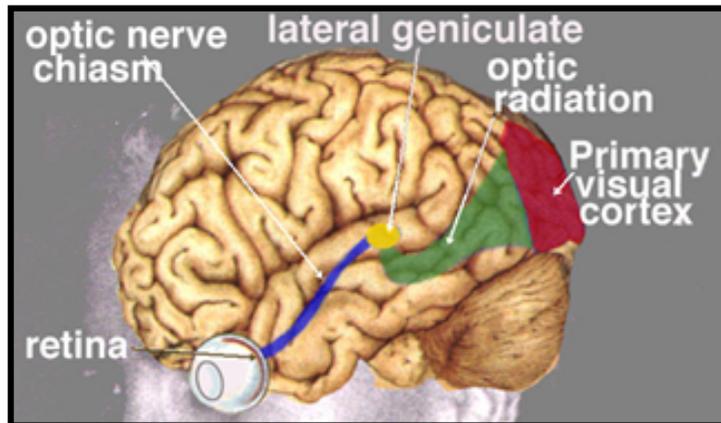
# Séance 5: Rendu et perception

- Rendu et perception en synthèse d'image
  - intro au système visuel humain et ses limites
  - contrast sensitivity function, visible difference predictors, change blindness, etc.
  - Applications:
    - compression d'images/vidéo
    - « tone mapping » et affichage
    - gestion de niveaux de détail géométrique et textures
    - éclairage
    - comparaison d'images / évaluation de qualité

# Séance 5: Rendu et perception

- Rendu et perception en acoustique
  - intro au système auditif humain et ses limites
  - masquages, loudness, perception 3D, auditory scene analysis
  - Applications:
    - compression / évaluation de qualité
    - traitement du signal rapide
    - priorités de source sonores et « audibility culling »
    - spatialisation de scènes sonores complexes

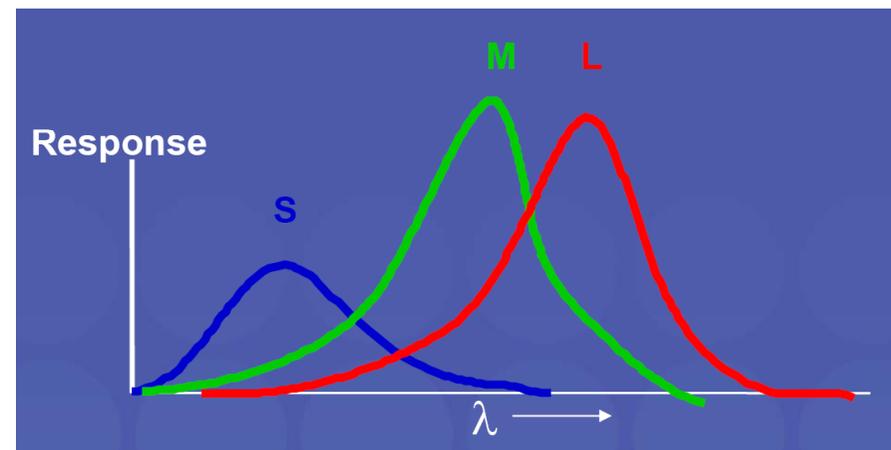
# Systeme visuel humain



- Structure physiologique établie
- Comportement perceptif est très complexe

# La rétine

- Convertit la lumière en signaux vers le cerveau
- Batonets
  - achromatiques
  - seulement sensibles aux basses intensités
- Cônes
  - vision en couleurs
  - sensibles seulement aux hautes intensités
  - 3 types



# Cones vs. Batonets



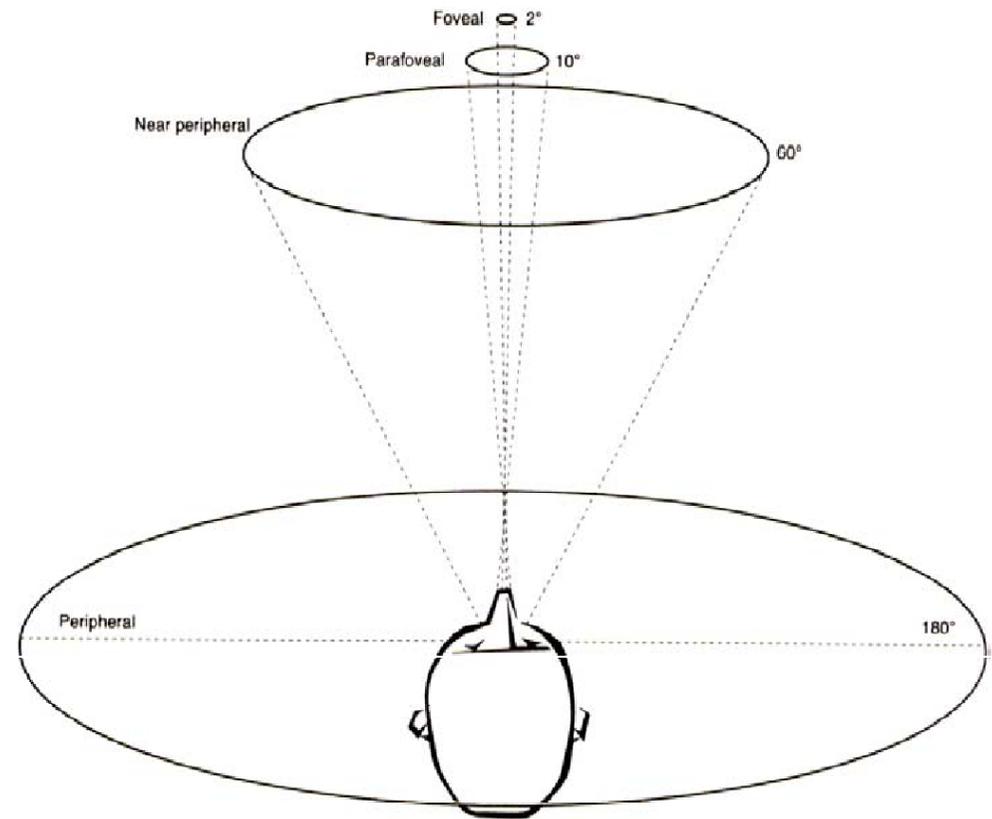
**A Multiscale Model of Adaptation and Spatial Vision for Realistic Image Display.  
Pattanaik, Ferwarda, Fairchild, Greenberg, SIGGRAPH98**

# Caractéristiques de la vision humaine

- Variation de la sensibilité
  - adaptation visuelle à une large plage de dynamique
- Adaptation chromatique
  - estimation de la couleur de l'éclairage
  - abstraction de cette couleur dans la perception de la couleur "propre" des objets

# Visual environment exploration

- Acuity not uniform
- Align fovea with relevant features
- Explore our visual environment with gaze movements
- How we stitch all these observations together is still a mystery

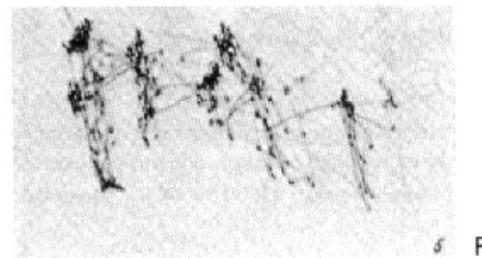
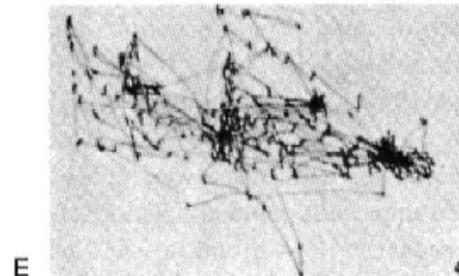
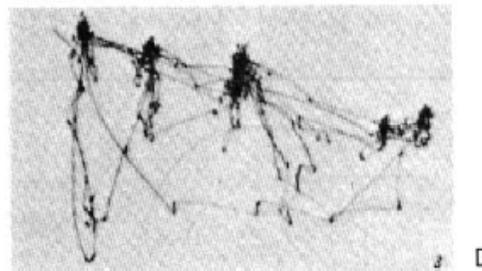
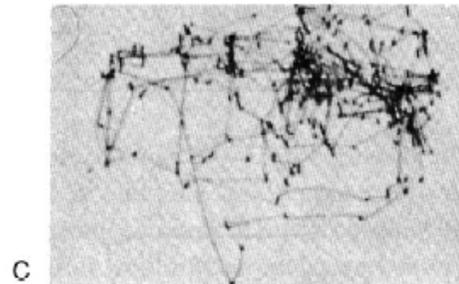
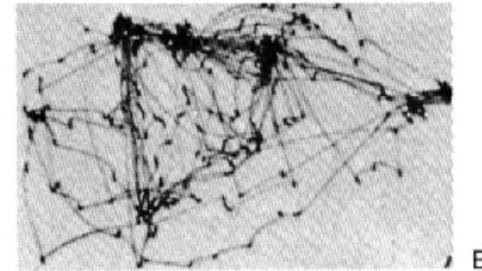


# Saccade

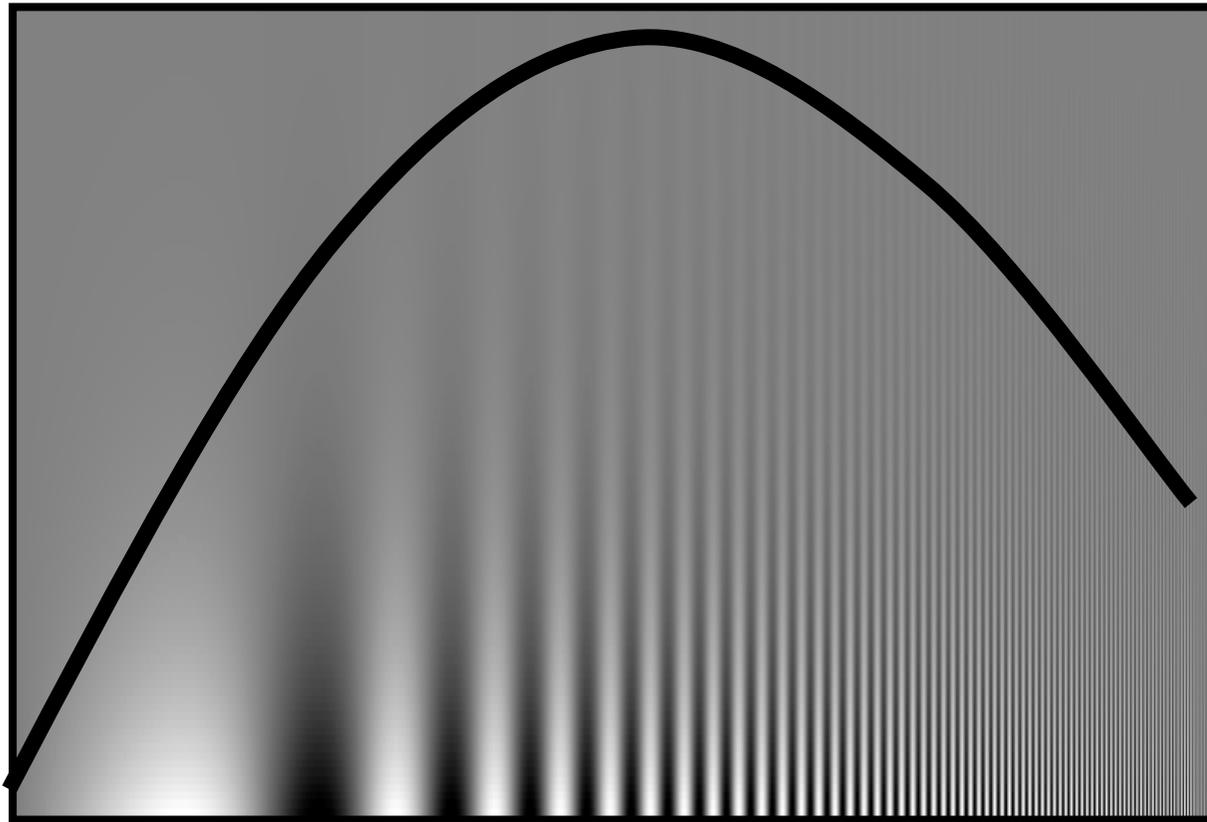
- Used to scan the visual field
  - Can be controlled
- Two phases
  - Ballistic movement: 30 ms and up to  $900^\circ/s$
  - Fixation  $\sim 300ms$
- Saccadic suppression
  - No blur is experienced during the ballistic movement
  - We “suppress” our vision while the gaze moves

# Dépend de la tâche

- painting by Repin
- B: free
- C: economic level
- D: ages
- E: what were they doing
- F: remember cloth



# Contrast sensitivity

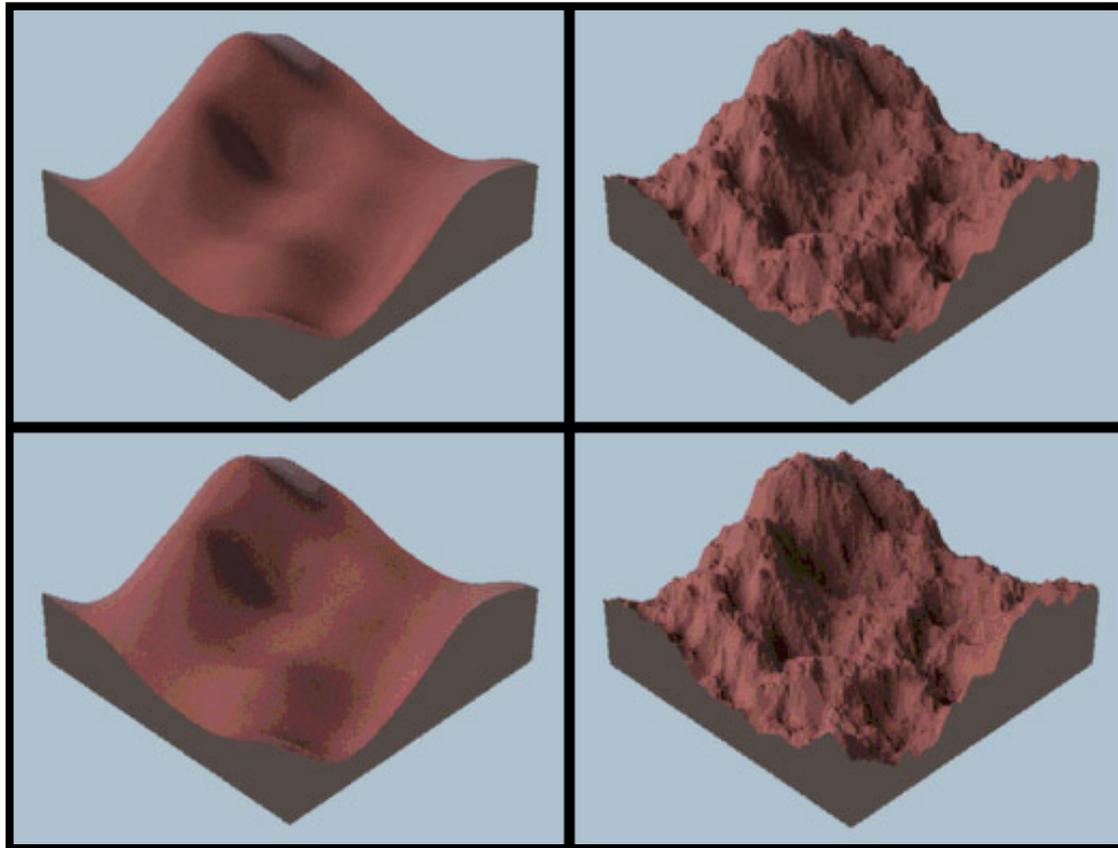


Campbell-R obs on contrast sensitivity chart

# Masquages visuels

- "Visual masking" is the phenomenon in which one visual pattern affects the detectability of another.

# Visual masking

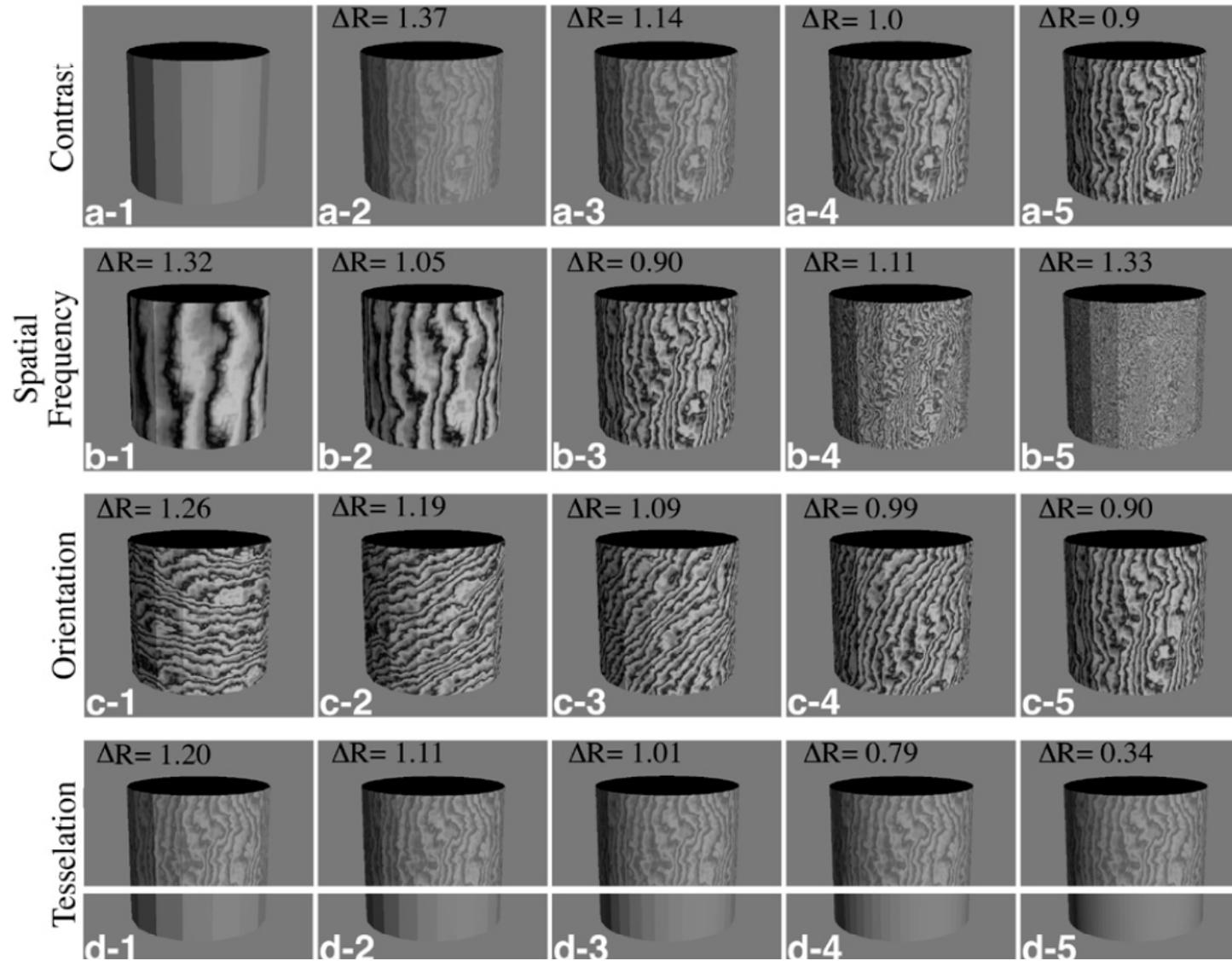


**A Frequency based ray-tracer.  
Bolin & Meyer, SIGGRAPH95**

# Visual Difference Predictors

- Probabilité par pixel qu'un utilisateur détecte une différence entre 2 images [Daly93]
- En général comprend 3 modèles:
  - sensibilité à la luminance  
(TVI threshold vs. intensity)
  - sensibilité au contraste  
(CSF contrast sensitivity function)
  - masquage visuel

# Visual Difference Predictors



# Visual Difference Predictors

- Trop coûteux pour le temps réel
- Nécessité de développer des approximations

# “Change blindness”

- failures to notice large changes to scenes
  - perturbation de l’attention
- Also *inattentional blindness* — the failure to notice unusual and salient events in their visual world when attention is otherwise engaged and the events are unexpected

# “Change blindness”



# Compression d'images

- Stocker de l'information visuelle de manière compacte et la restituer sans différence visible
- Compression avec perte
  - standards jpeg, etc.

# Compression d'images

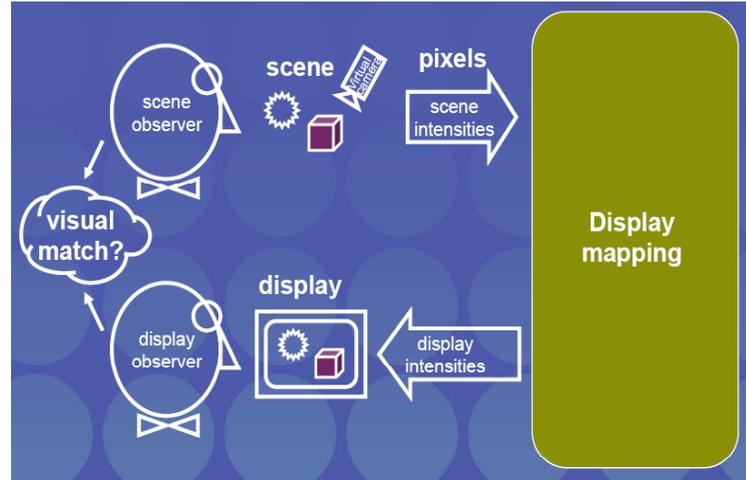
- Principe (e.g. JPEG):
  - décomposition de l'image en blocs 8x8
  - DCT
  - quantification variable par division des valeurs de la DCT par une matrice de quantification (transmise avec l'image)
  - arrondi à la valeur entière la plus proche

# Compression d'images

- Optimisation de la matrice de quantification
  - utilisation de visual difference predictors
  - produit une matrice optimale pour une image donnée (i.e., qui minimise les différences visibles)
  - e.g. DCTune de la NASA

# Affichage et “tone mapping”

- afficher de manière “fidèle” une image de synthèse
  - pb: dynamique du système d’affichage

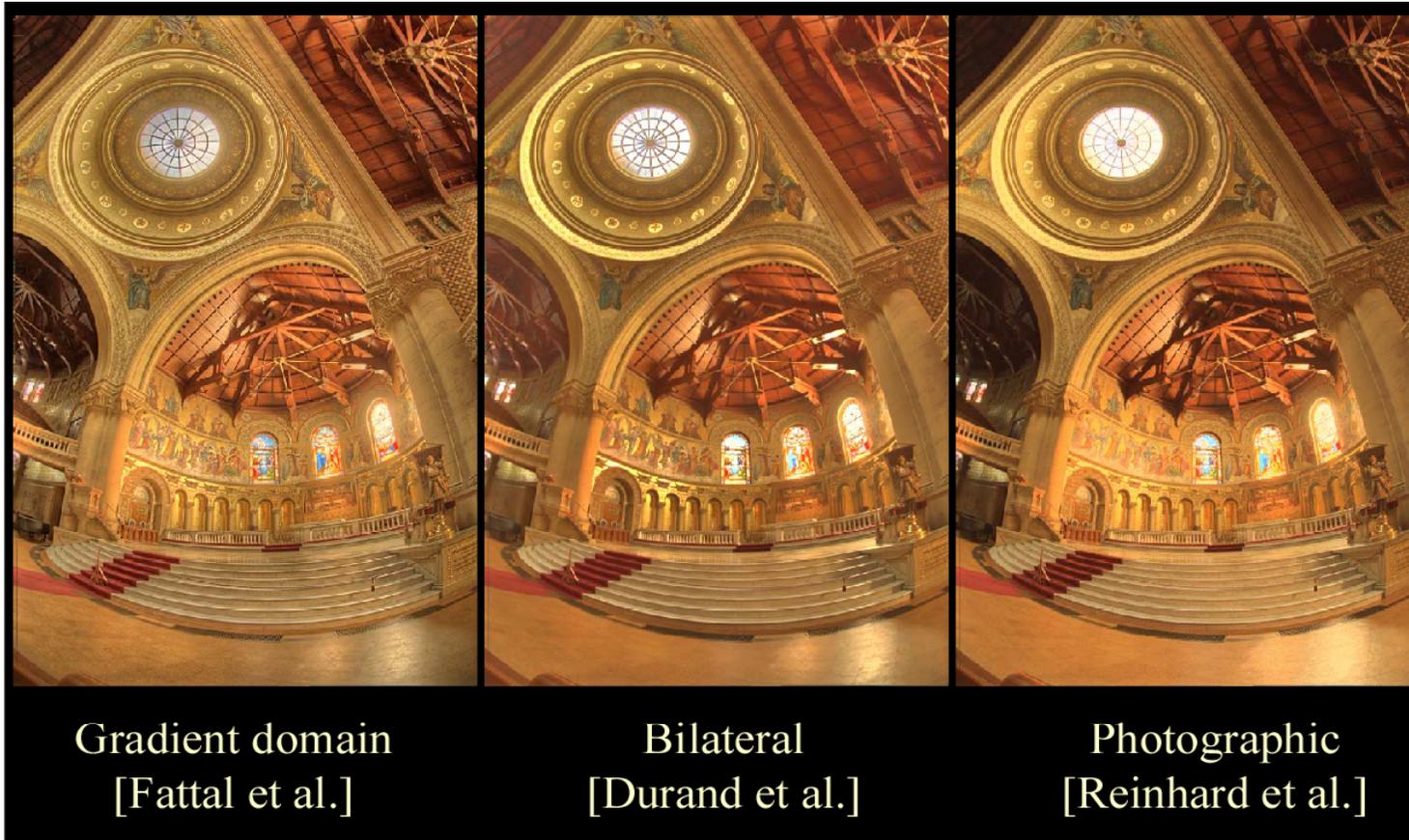


**Voir également : A review of tone reproduction techniques.  
K. Devlin, Tech. Report, Univ. Of Bristol, 2002 .**

# Affichage et “tone mapping”

- Gamma correction
  - relation entre le voltage d’entrée d’un écran et l’intensité lumineuse observée est non linéaire
  - $\text{new\_pixel\_value} = \text{old\_pixel\_value}^{(1.0/\text{gamma})}$
- More general operators
  - Based on image formation
  - Based on the human visual system
  - Global operators vs. Local operators
  - Time dependence

# Affichage et “tone mapping”



# Affichage et “tone mapping”



Gradient domain  
[Fattal et al.]

Bilateral  
[Durand et al.]

Photographic  
[Reinhard et al.]

# Gestion de niveaux de détail

- Choisir un niveau de détail minimal pour la géométrie et les textures sans différence visible
- e.g., le contenu d'une texture appliquée sur une surface peut masquer les artefacts visuels due à la simplification

**Image-driven simplification. Lindstrom & Turk.  
ACM Transaction on Graphics. 2000**

**Perceptually-guided simplification of Lit, Textured Meshes.  
Williams, Luebke, Cohen, Kelley, Schubert.  
ACM Symposium on Interactive 3D Graphics, 2003**

# Gestion de niveaux de détail

- Prédeterminer l'erreur perceptive de chaque simplification possible
  - stockée sous forme de graphe (multi-triangulation)
- Déterminer en temps-réel une coupe dans le graphe
  - minimize l'erreur en respectant un budget de triangles donné

# Simulation de l'éclairage

- Effectuer la simulation d'éclairage en ne traitant les phénomènes proportionnellement à leur importance visuelle
- Eliminer les composants de l'éclairage "invisibles" (e.g., ombres)
  - Just Noticeable Differences
- Rendu en domaine fréquentiel

# Simulation de l'éclairage

- Déterminer comment échantillonner l'éclairage indirect pour l'illumination globale

A Perceptually-based physical error metric for realistic image synthesis .  
Ramasubramanian, Pattanaik, Greenberg, SIGGRAPH99



(a) reference global illumination solution



(b) sample density pattern



(c) adaptive global illumination solution

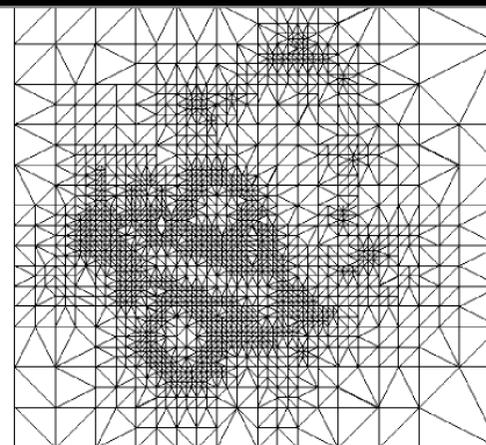
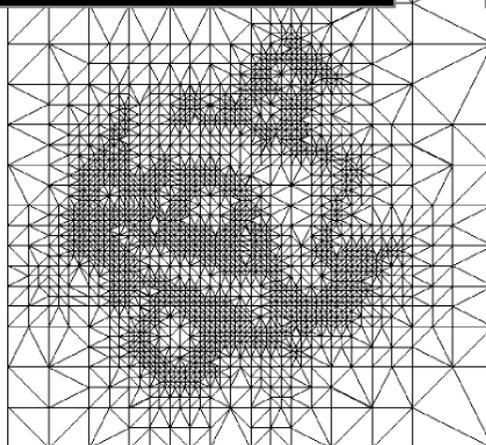
# Radiosité et ombres



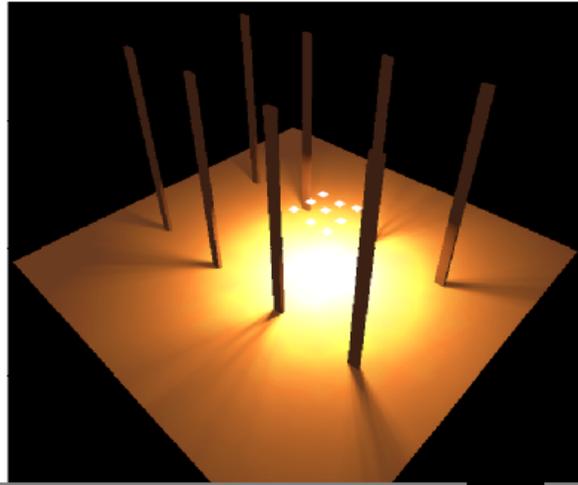
Input mes h



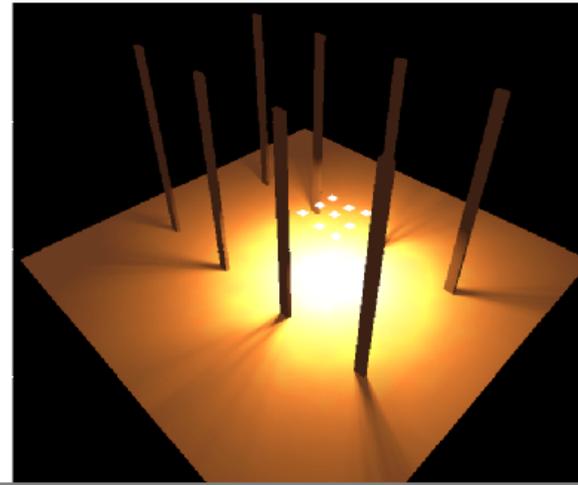
Optimis ed mes h



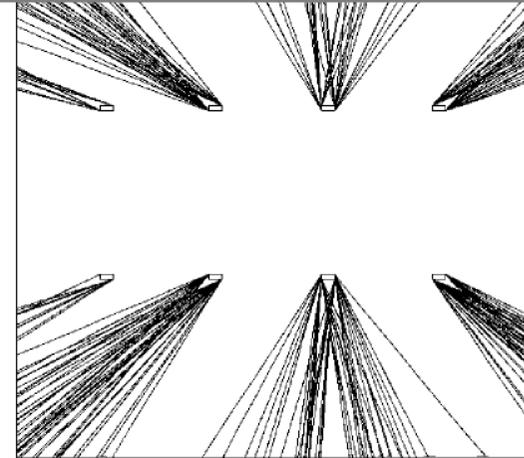
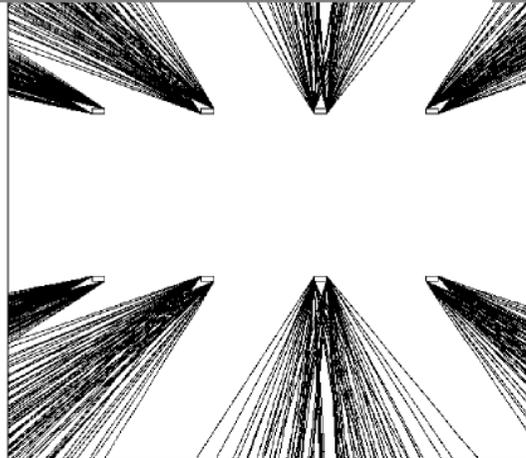
# Radiosité et ombres



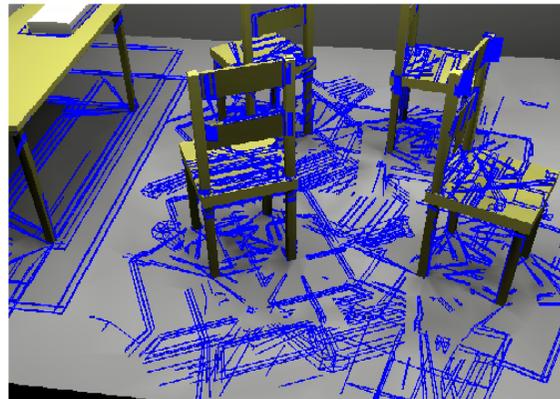
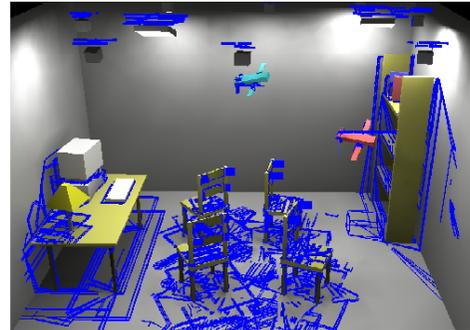
Original scene



Culled discontinuities



# Radiosité et ombres



**Fast and Accurate Hierarchical Radiosity Using Global Visibility**  
Frédo Durand, George Drettakis and Claude Puech.  
ACM Transactions on Graphics, April 99

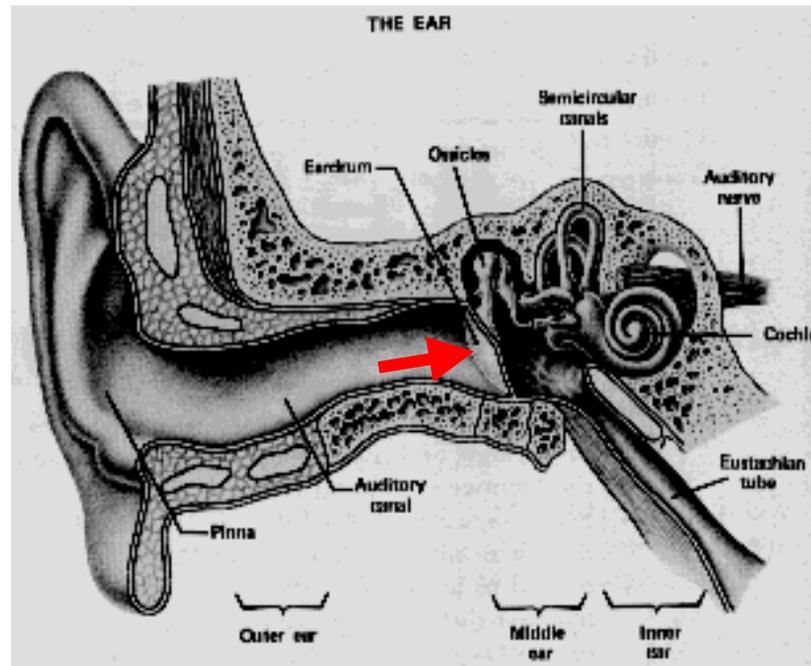
# Références

- Perceptual and Artistic Principles for Effective Computer Depiction (Frédo Durand)
  - [http://graphics.csail.mit.edu/~fredo/SIG02\\_ArtScience](http://graphics.csail.mit.edu/~fredo/SIG02_ArtScience)
- Change blindness
  - <http://viscog.beckman.uiuc.edu/change/>
  - <http://www.usd.edu/psyc301/ChangeBlindness.htm>
- Level of Detail for 3D Graphics
  - D. Luebke, M. Reddy, J. Cohen, A. Varshney, B. Watson, and R. Huebner; Morgan Kaufmann. ISBN 1-55860-838-9.

# Références

- James Ferwerda (Cornell)
  - <http://www.graphics.cornell.edu/~jaf/>
- Bolin & Meyer
  - <http://www-users.cs.umn.edu/~meyer/publications.html>
- Vision Science, from photons to phenomenology
  - Stephen E. Palmer. Bradford Books.

# Systeme auditif humain



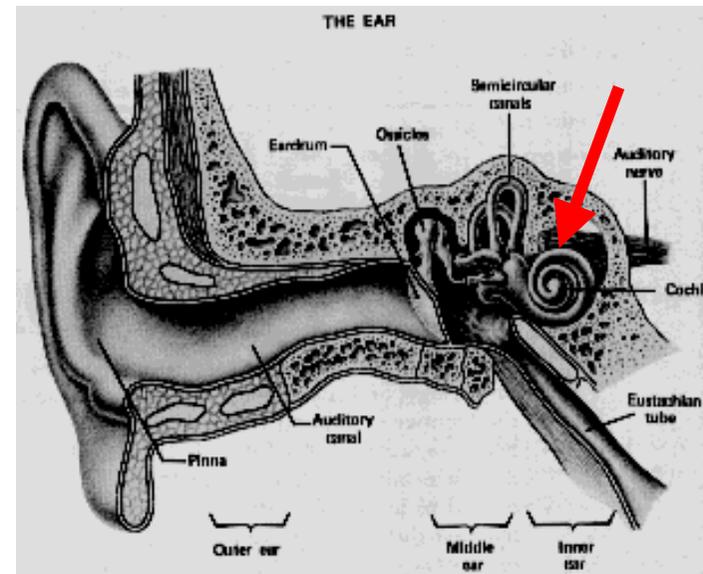
- L'oreille est un récepteur complexe
  - le tympan mesure la pression acoustique
  - convertie en atomes temps-fréquence par le cortex

# Perception du son

- Fréquences audibles: 20Hz- 20kHz
- Dynamique:  $10^{12}$  W.m<sup>2</sup> (120 dB) du seuil d'audition au seuil de douleur

# Perception du son

- Filtrage en « **bandes critiques** »
  - membrane basilaire (cochlée)
  - le son est découpé en **25 bandes** de fréquences
  - 1/3 octave



# Masquage

- **Effet de précedence** : pour des signaux similaires le premier arrivé à l'auditeur masque le second (essayez avec votre chaine hifi)
- **Masquage temporel** : un son fort masque un son faible immédiatement après...et même avant !
- **Masquage fréquentiel**: une fréquence masque les fréquences voisines (+ dans les basses)

# Compression audio perceptive

- Stocker de l'information sonore de manière compacte et la restituer sans différence audible
- Compression avec perte (inaudible)
  - standards MPEG I-layer III (*mp3*), MPEGII AAC, DolbyDigital/DTS, etc.

# Compression audio perceptive

- Principe:
  - Découper le signal en bandes critiques (Bark)
  - Allouer le nombre de bits de quantification dynamiquement
  - Mais réduire ce nombre introduit du bruit
  - Exploiter les masquages auditifs pour assurer que le signal masque le bruit de quantification

# Priorité des sources et “audibility culling”

- Application a la spatialisation du son
  - éviter de traiter des sources sonores inutilement
- Insérer progressivement les sources jusqu’à ce que leur combinaison masque les sources restantes
  - pré-calculs d’indicateurs sur les signaux
  - heuristique de priorité pour les sources

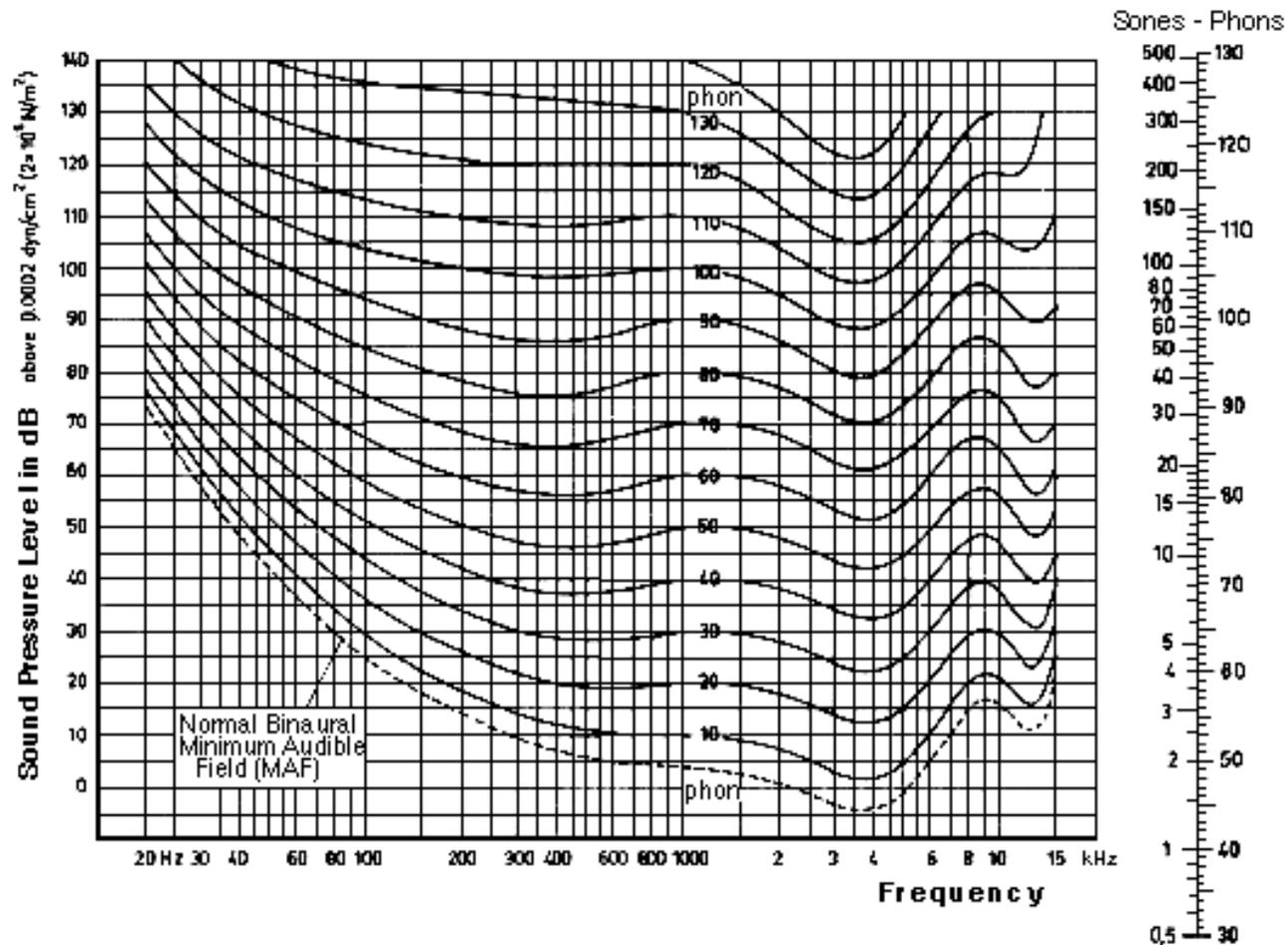
# Indicateurs du signal

- Spectre de puissance (bandes de fréquences)
- Tonalité (bandes de fréquences)
  - 0 = tone, 1 = bruit
  - les seuils de masquages en dépendent
- beaucoup d'autres utilisés, par exemple, pour le "music retrieval" dans des bases de données
- données compactes (calculées pour des blocs du signal d'entrée).

# Priorité des sources sonores

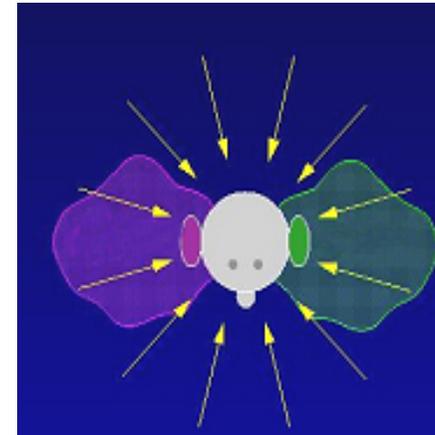
- Sonie (“loudness”)
  - modèle perceptif de l’intensité perçue des sons
  - dépend du type de son (noise/tone/...), de la fréquence et du niveau de pression
  - unité : le phon
- Calculée au run-time à partir du spectre de puissance
  - utilisée pour trier les sources

# Equal loudness contours of pure tones



# Perception 3D du son

- la perception **binaurale** (des 2 oreilles) donne des indications de **direction de provenance en 3D**



- différence de temps d'arrivée aux deux oreilles
- interaction du son avec la tête et le buste
  - filtrage directionnel

# Limites de la perception audio 3D

- Résolution spatiale non-uniforme
  - faible dans le plan vertical, élevée dans le plan horizontal
  - faible à l'arrière, élevée à l'avant
- Résolution limitée pour les sources en mouvement

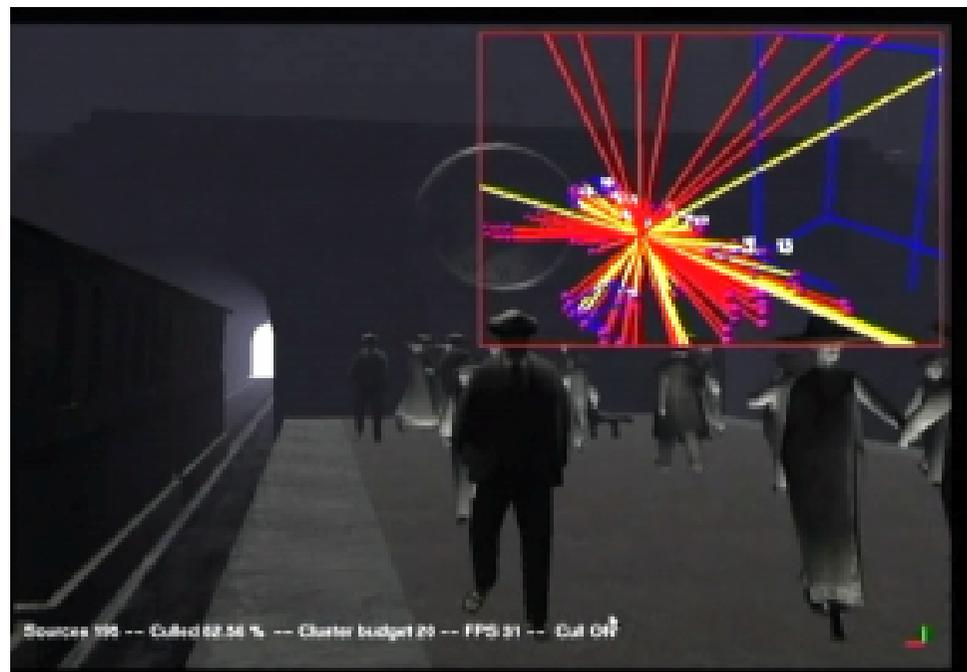
# Groupement de sources pour le rendu audio 3D

- Pb: nombre limité de canaux audio hardware ou faible puissance CPU
  - factoriser les traitements de spatialisation pour en amortir le coût

# Groupement de sources pour le rendu audio 3D

- Pb: nombre limité de canaux audio hardware ou faible puissance CPU
  - factoriser les traitements de spatialisation pour en amortir le coût
- Solution : grouper les sources spatialement proches
  - définir une position représentative pour le groupe
  - barycentre pondéré par l'importance

# Démo



# Références

- Introduction to the psychology of hearing
  - Moore. Academic Press
- Applications of DSP to audio and acoustics
  - Kahrs/Brandenburg. Kluwer Academic Publishers
- <http://www.mp3-tech.org/>
- Chez nous
  - [www-sop.inria.fr/reves/](http://www-sop.inria.fr/reves/)

# Conclusions

- Vision et audition ont des limites
- Physiologie vs. cognition
- On retrouve les mêmes phénomènes
  - e.g., masquages, “fovea”
  - exploités pour la compression
- Leur exploitation offre de nombreuses perspectives pour le rendu audio et visuel
  - qualité de reproduction
  - efficacité des simulations