

Entre perception et images numériques : une perspective.

Samuel Delepouille - habilitation à diriger des recherches.

jury :

- Pr Rémi Cozot (président)
- Dr Solène Kalenine (examinatrice)
- Pr Céline Loscos (rapporteure)
- Pr Laurent Madelain (examinateur)
- Pr Pascal Mamassian (rapporteur)
- Pr Daniel Meneveaux (rapporteur)
- Pr Christophe Renaud (garant)

Plan

- 1. Problématique générale**
2. Rendu photoréaliste
3. Nos travaux
4. Perspective / Conclusion

Problématique générale

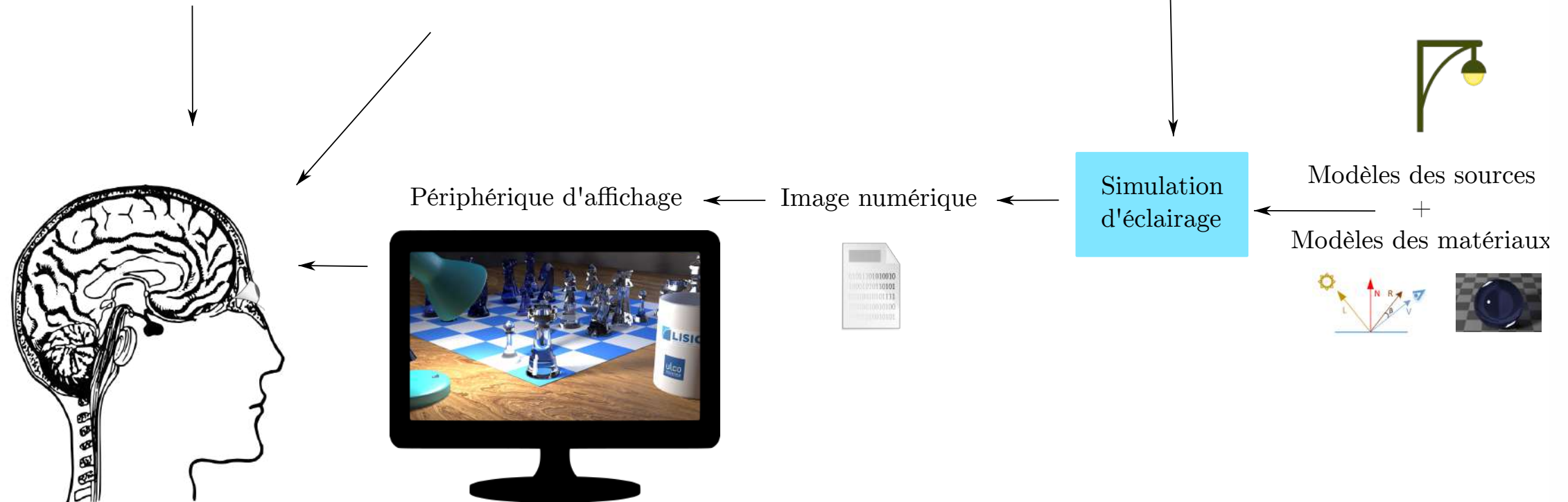
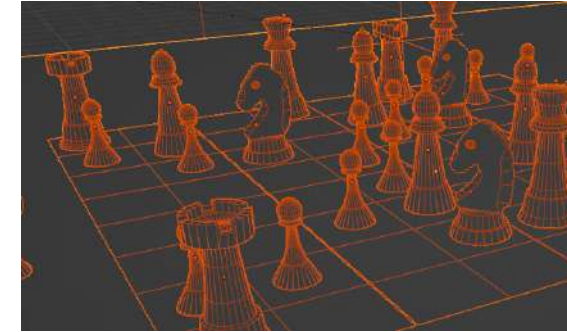
peinture



Scène réelle

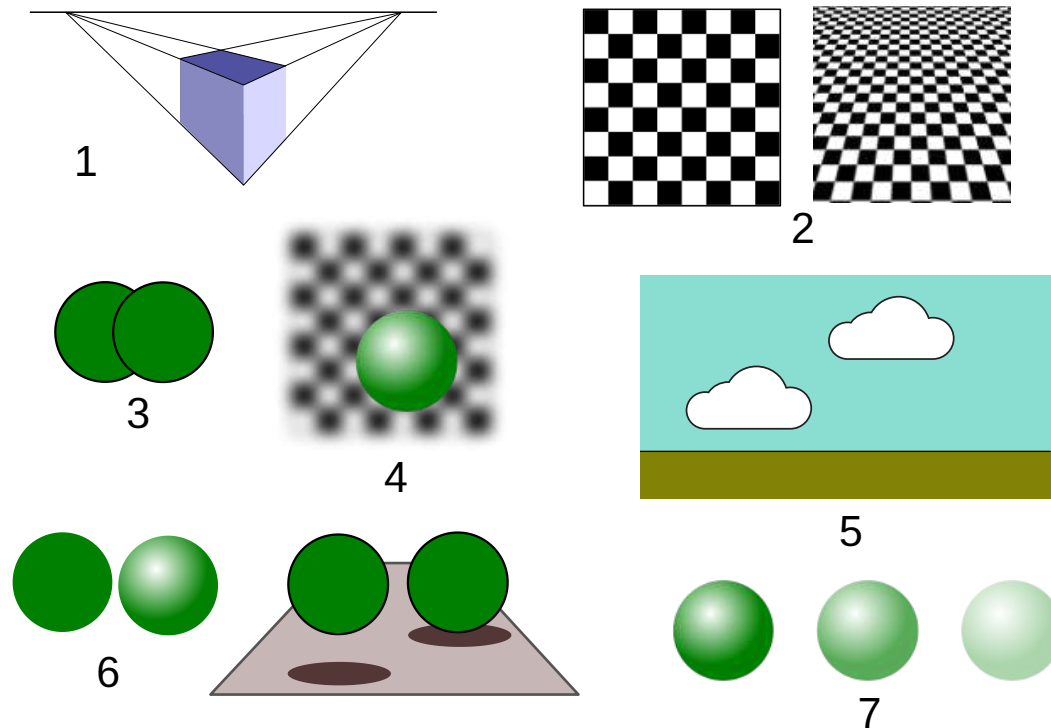


Modèle géométrique

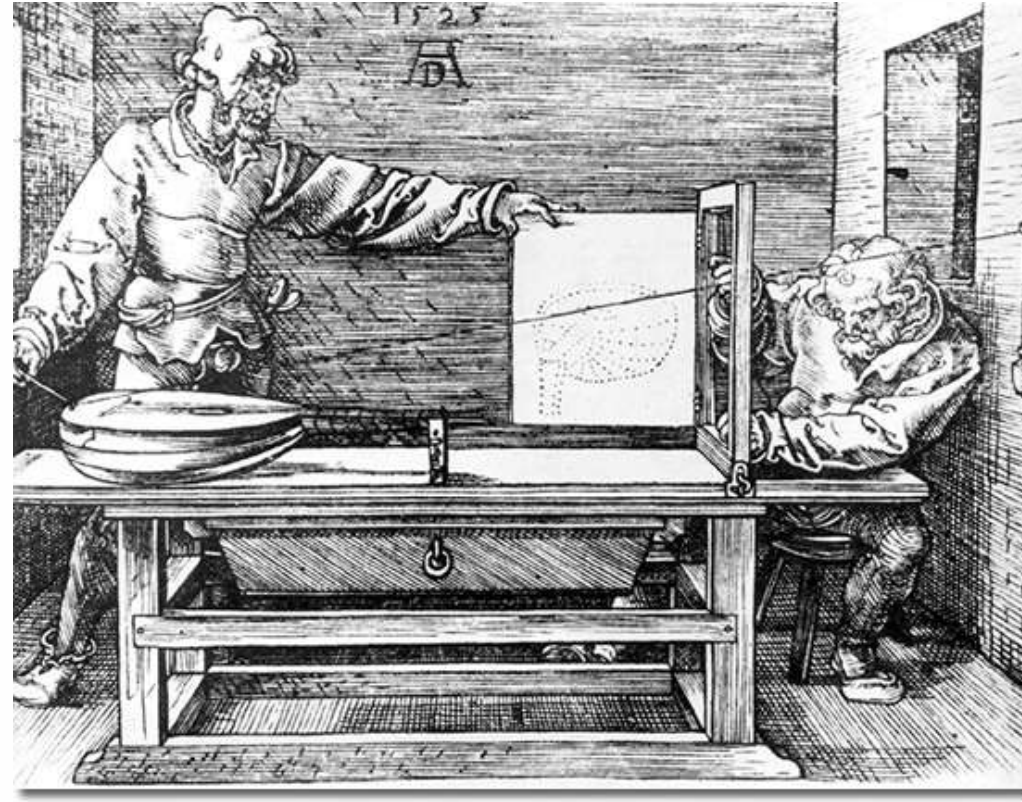


Indices de perception de la distance

- Comment provoquer une impression de 3D à partir d'une image 2D ?
- Principes connus des peintres.
- Étudiés par la psychologie de la perception.
- Exploités par l'image numérique.



Fabrique de l'image



Albrecht Durer au 16e siècle ;

Actuellement : plusieurs millions par seconde mais même principe.

Plan

1. Problématique générale
2. **Rendu photoréaliste**
3. Nos travaux
4. Perspective / Conclusion

Techniques d'imagerie numérique

	Temps Réel	Photoréalisme
Temps de rendu	< 40 ms : 25 images par seconde	minutes, heures, jours
Techniques	Illumination locale, GPU	Illumination globale
Applications	Jeux vidéo, simulation, Réalité virtuelle & augmentée	Conception d'objet, Cinéma, art, architecture, histoire

Rendu photoréaliste : pourquoi ?



DANTHREE



Applications à la reconstitution historique



reconstitution virtuelle du Pont Notre Dame (18e siècle)

Applications à la perception du mouvement

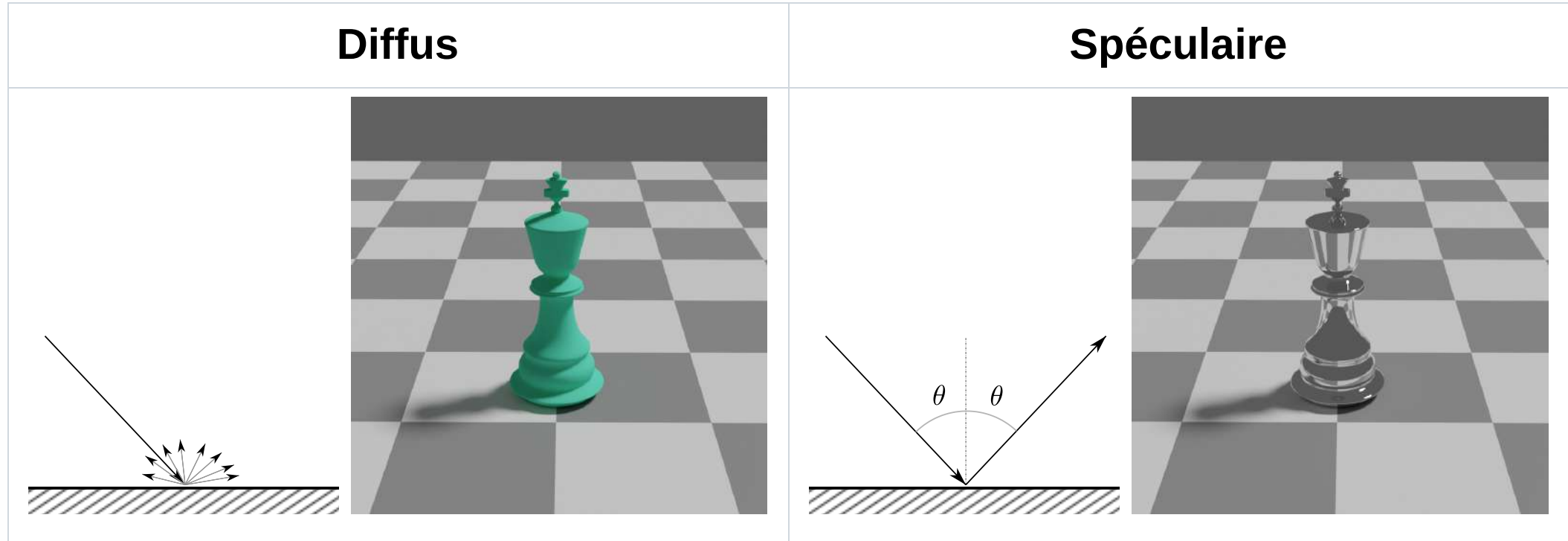


conflictual

non conflictual

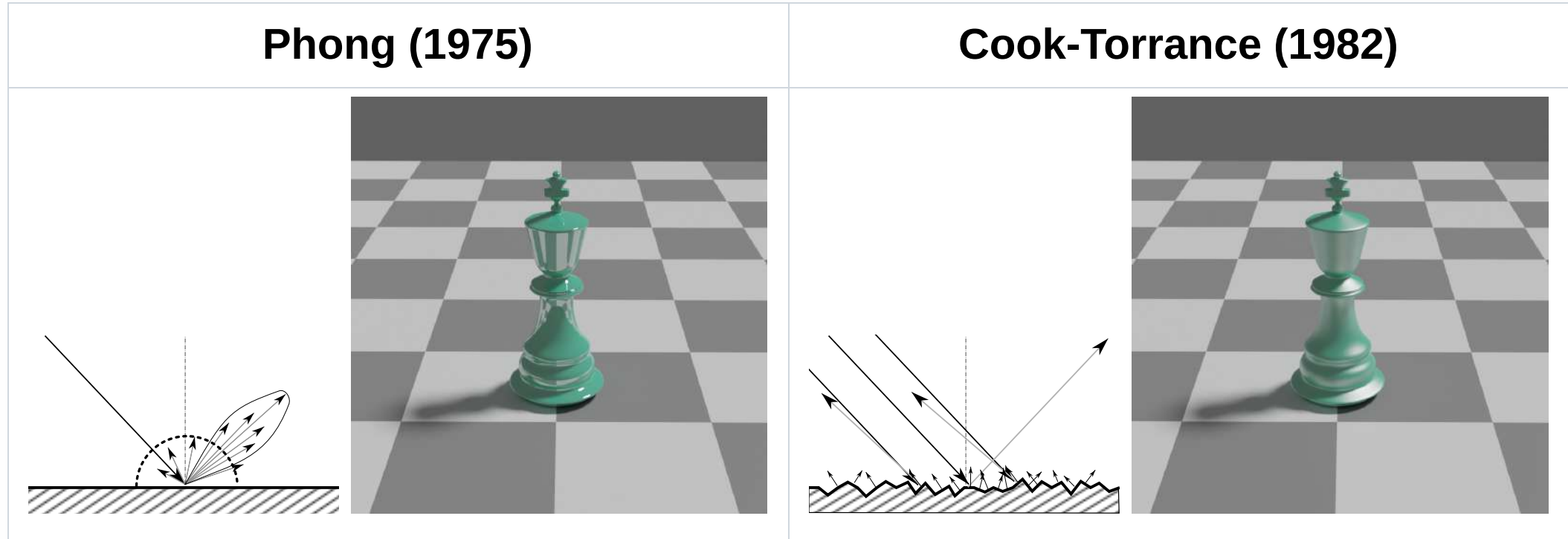
Rendu photoréaliste : comment ?

Modèles de matériaux : modèles de BRDF



Rendu photoréaliste : comment ?

Modèles de matériaux : modèles de BRDF



Equation du rendu

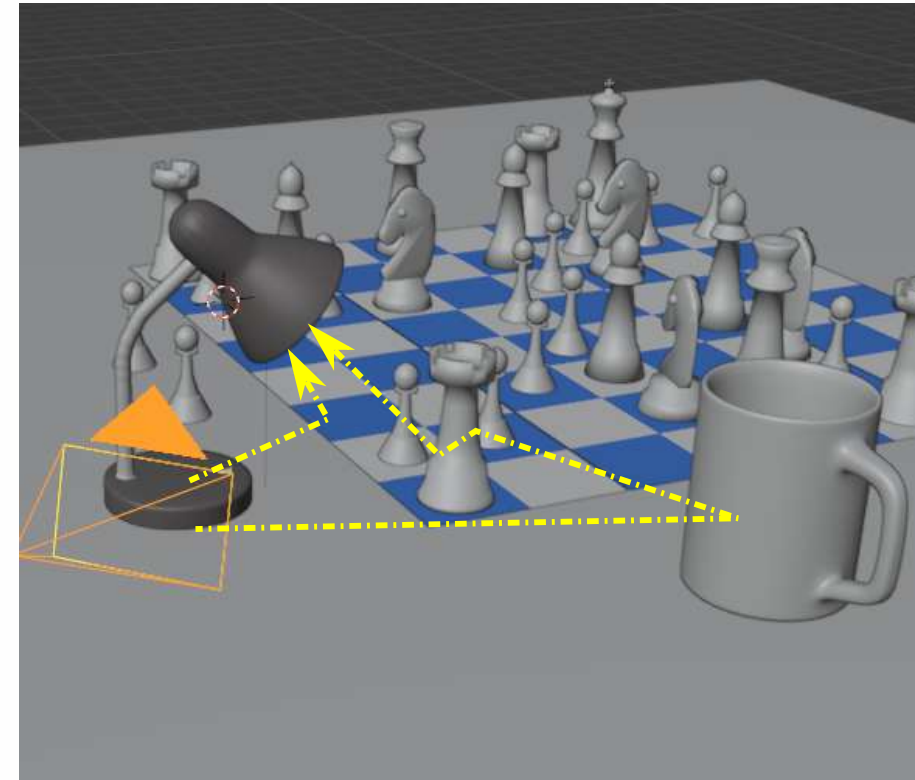
(Kajiya, 1986)

$$L_r(x, \vec{\omega}_r) = L_e(x, \vec{\omega}_r) + \int_{\Omega} f_r(x, \vec{\omega}, \vec{\omega}_i) \cdot L_i(x, \vec{\omega}_i) \cdot \cos(\vec{N}_x, \vec{\omega}_i) d\vec{\omega}_i$$

Tracé de chemins méthode de Monte-Carlo (version de base)

1. Choisir un rayon pour un pixel (poids = 1)
2. Trouver son intersection avec la surface la plus proche
3. Tirer direction de réflexion (fonction du matériau), poids *= reflectance et aller en 2.

fin (roulette russe) : retourner poids * L_e
= calcul d'un échantillon \Rightarrow moyenne



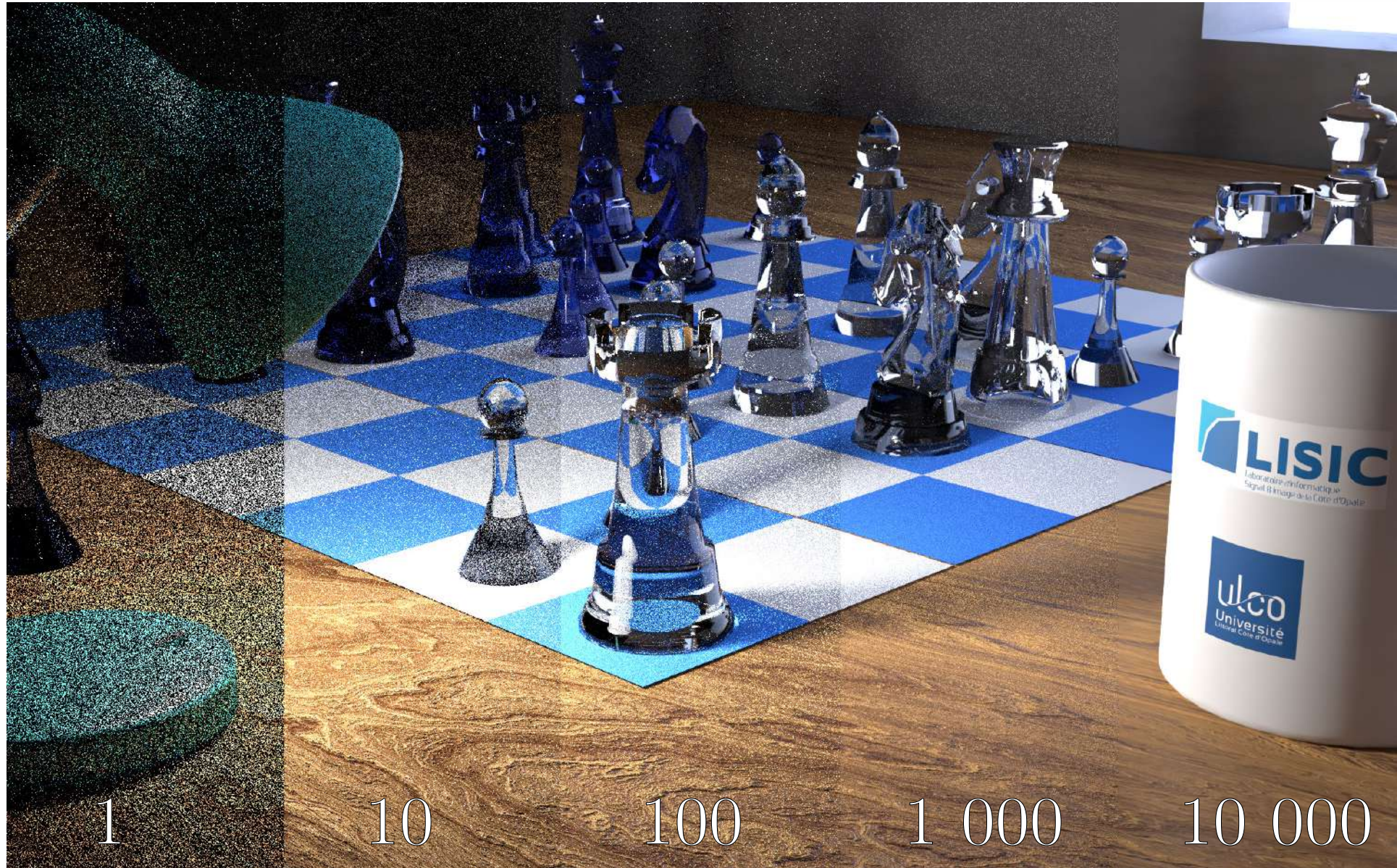
Avantages

- tous types de géométries (polygones, courbes, procédurales, milieux participants ...)
- tous types de BRDF (spéculaire, diffus, ...)
- économique en mémoire
- calcul au niveau de chaque pixel = facilement parallélisable
- méthodes non biaisées = l'erreur est un bruit dans l'image finale

Inconvénients

- convergence lente
- bruit résiduel dans l'image

Bruit en fonction du nombre d'échantillons





Bruit visuel

1. fortement perceptible
2. haute fréquence spatiale
3. proche d'un bruit « poivre et sel »
4. répartition hétérogène et peu prédictible
5. sa quantité diminue avec le nombre d'échantillons
(dans les bons cas)

Problème : quand arrêter le calcul ?

- bruit = un artefact visuel
 - impose un critère humain vs numérique
- problème de seuils perceptifs
 - si $<$ au seuil = image de mauvaise qualité
 - si $>$ au seuil = temps de calcul inutile
- n problèmes de seuils avec n le nombre de pixels dans l'image.
- image full HD = 2 M de seuils !

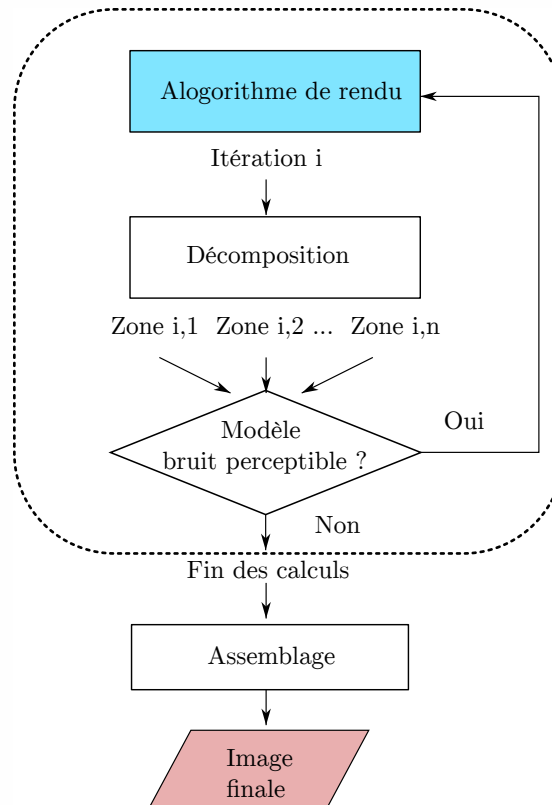
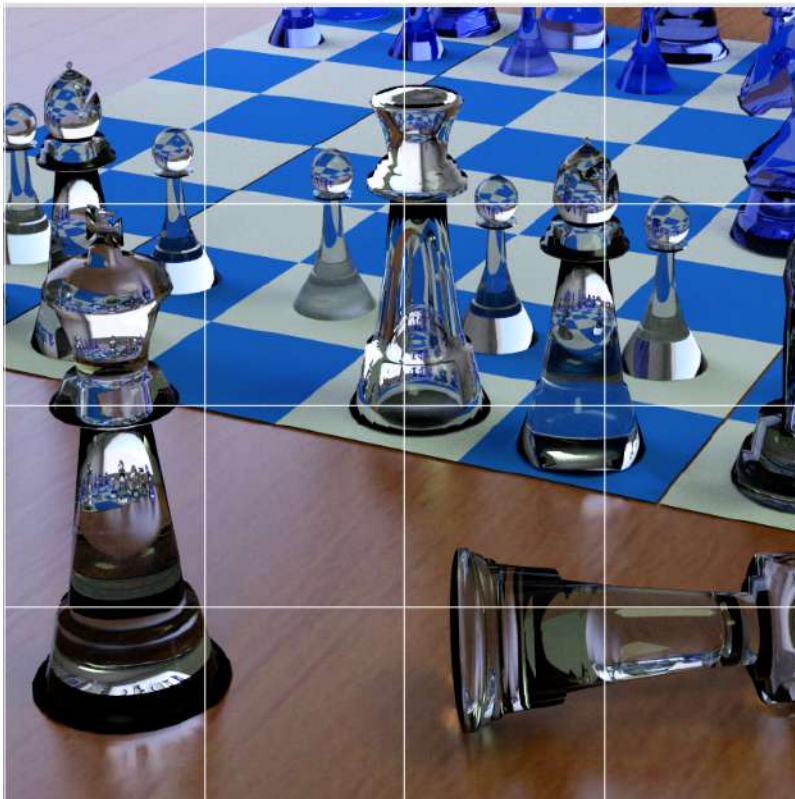
Plan

1. Problématique générale
2. Rendu photoréaliste
3. **Nos travaux**
4. Perspective / Conclusion

Comment étudier le bruit ?

Thèse de Nawel Takouachet

- Sur toute l'image ? Pas assez précis. Au niveau du pixel ? Très long ...
- Découpage de l'image en zones rectangulaires.



Études préalables



10 000 spp



1020 spp

Études préalables



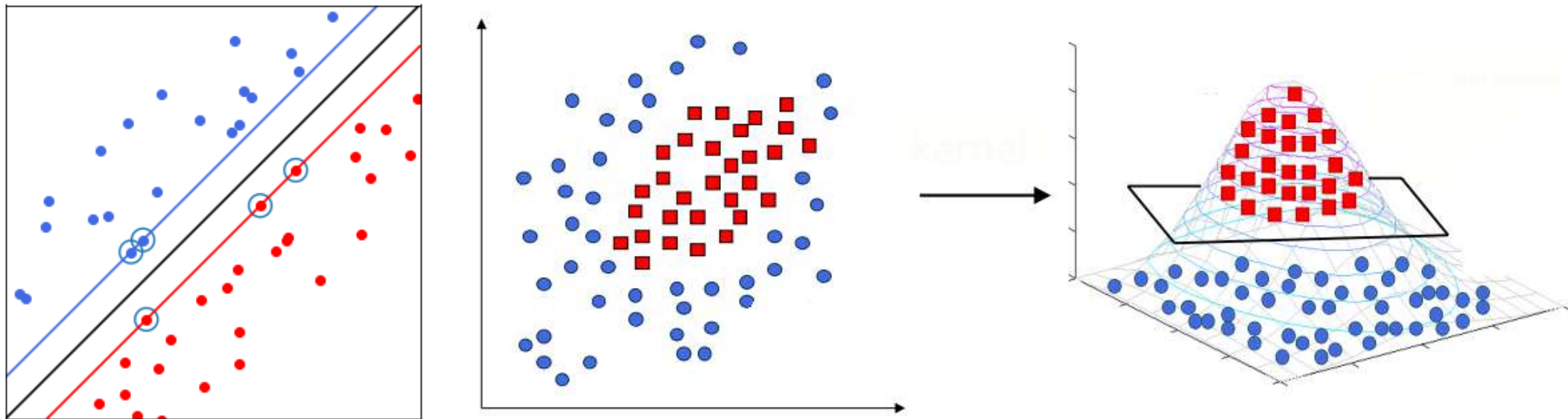
10 000 spp



1020 spp

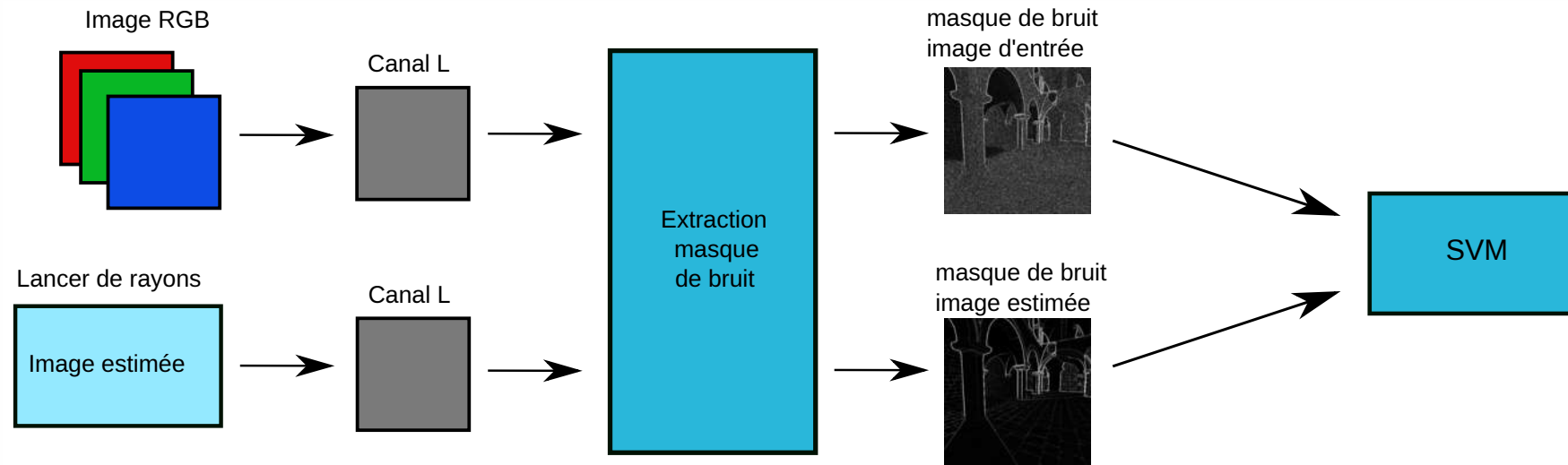
Peut-on apprendre ces seuils ?

- Décider si une image contient du bruit perceptible = problème de classification
- Première étude : SVM
- SVM = *Support Vector Machine* ou Machine à vecteurs de support
- Recherche de l'hyper-plan séparateur



Résultats

- Premiers tests : base de 16 images calculées par moteur PTWNEE
- Apprentissage SVM light
- échec
 - contenu > bruit
- guider l'apprentissage



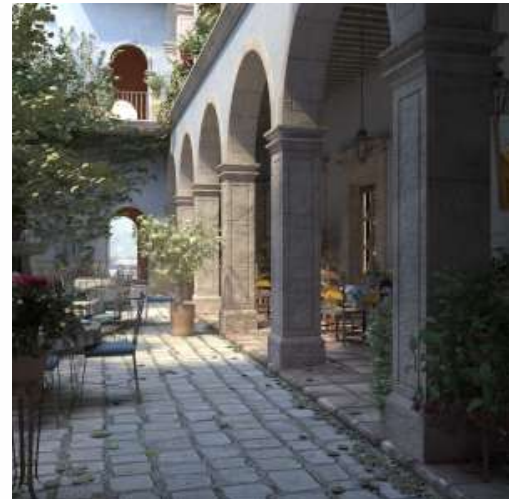
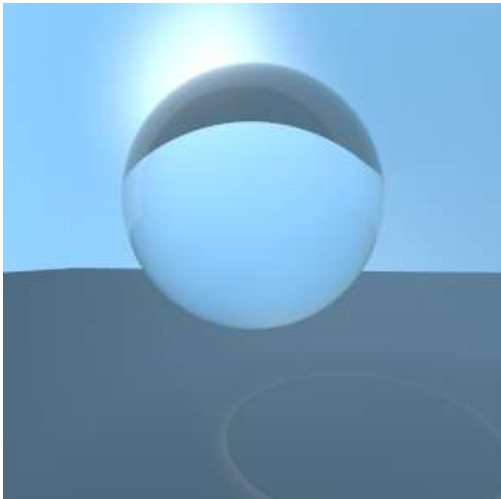
Comment aller plus loin ?

Thèse de Jérôme Buisine

- « Méthodes d'apprentissage automatique pour la *prise* en compte du bruit dans les images de synthèse » (soutenue le 8 décembre 2021)
- s'affranchir des limitations précédentes
- Augmenter la généralisation
 - variété et complexité des scènes
 - variété des effets lumineux
 - constitution d'une nouvelle base d'images
- Utiliser l'apprentissage profond
- Financement par l'ANR du projet PrISE-3D : Perception, Interactions et Simulation d'Éclairage 3D en collaboration avec SCALab (Thèse de Vasiliki Myrodia)

Nouvelle base d'images

- niveaux de bruit : de 20 à 10 000 échantillons par pas de 20
- 16 blocs de 200×200
- 80 images de taille 800×800 issues de 28 scènes 3D
- plus de 162 M de pixel
- publiée sur <https://zenodo.org/record/4964303>

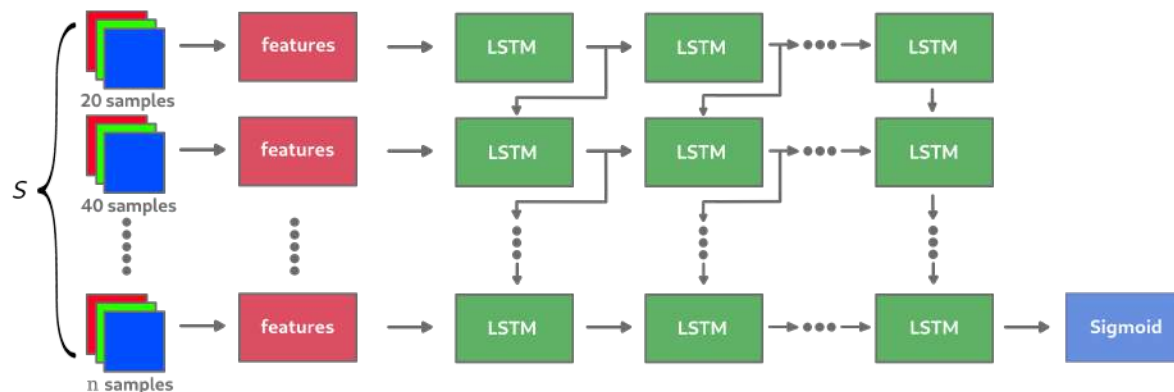


Approches testées (1)

- RNN : *Recurrent Neural Network*
 - utilisation d'une fenêtre glissante sur S images
 - SVD entropy

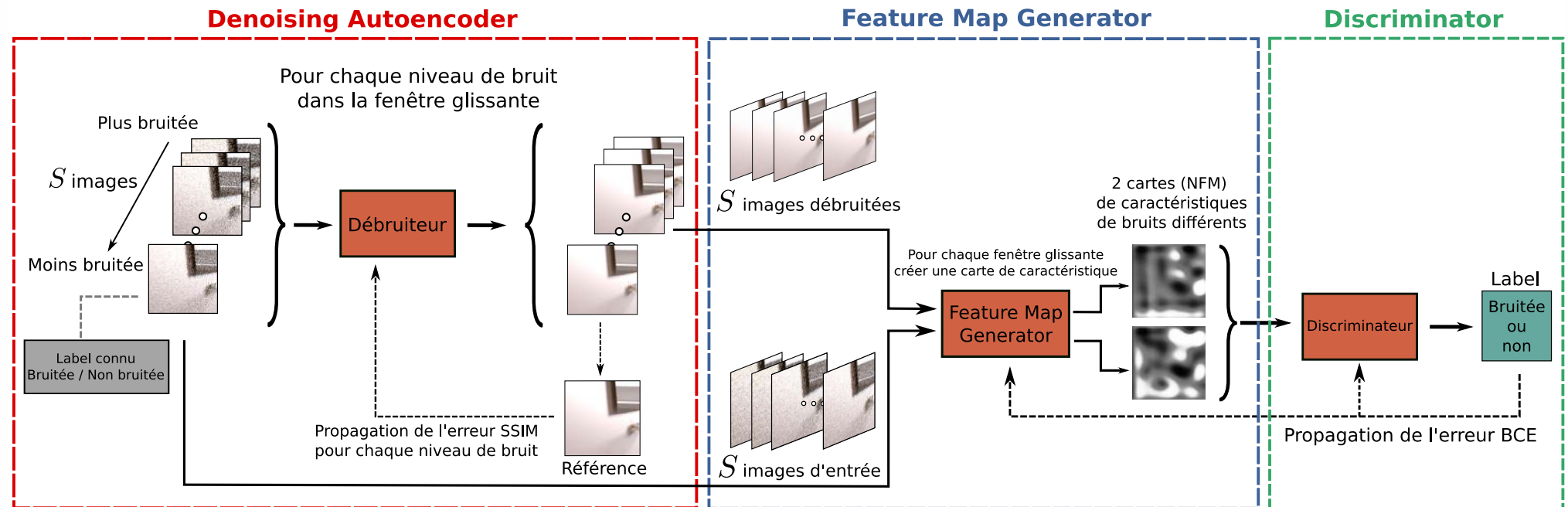
$$L = U \cdot \Sigma \cdot V^T$$

$$H = - \sum_{i=1}^M \bar{\sigma}_i \log_2(\bar{\sigma}_i)$$



Approches testées (2)

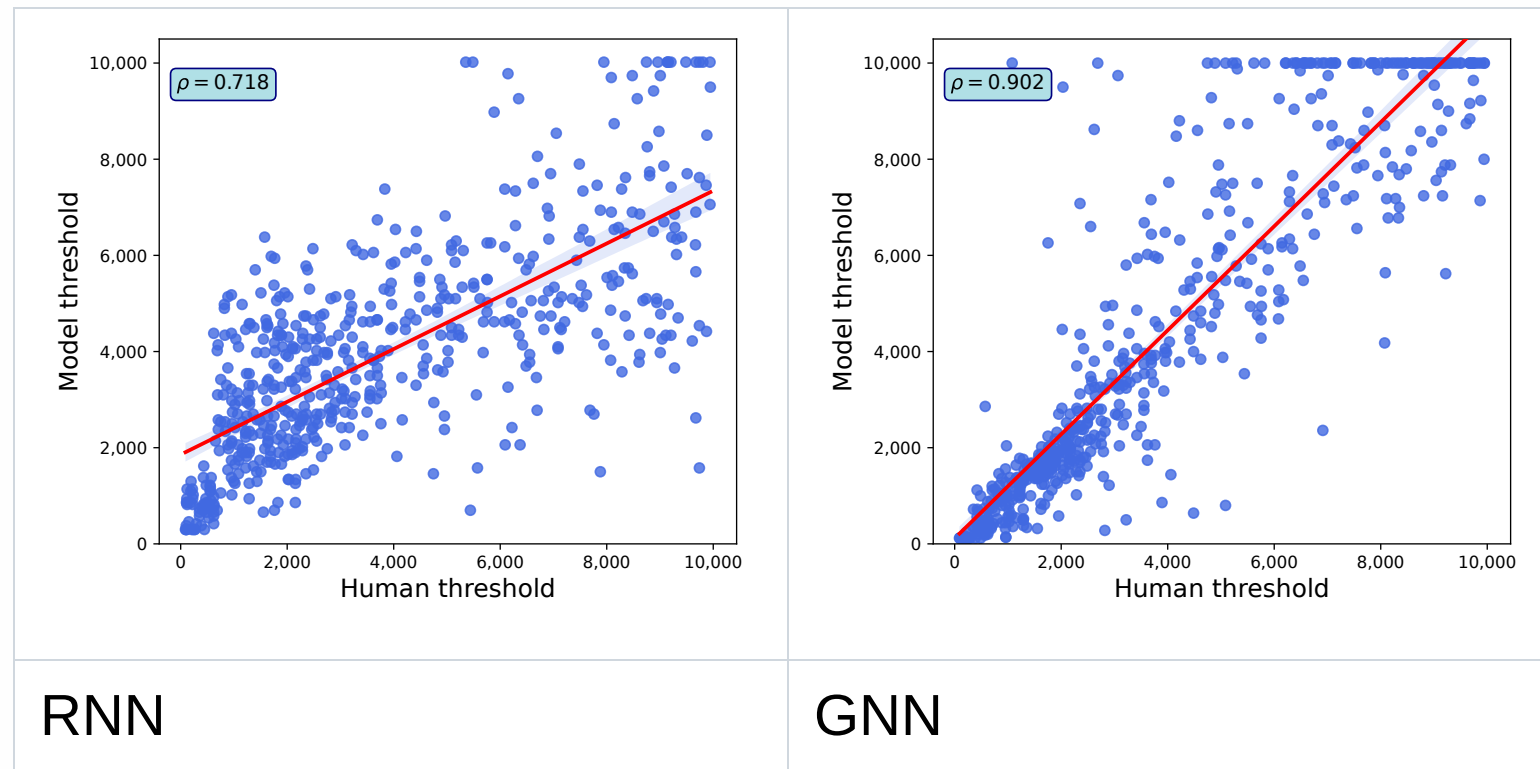
- GGN : *Guided-Generative Network*
 - rechercher une extraction automatique des caractéristiques du bruit.



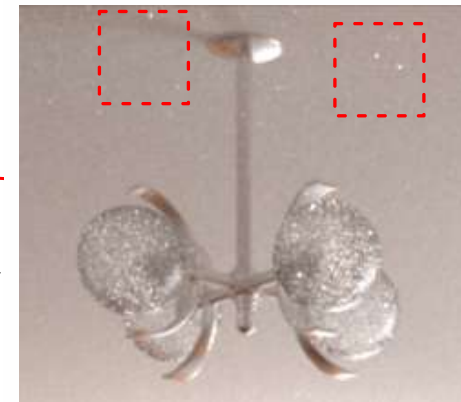
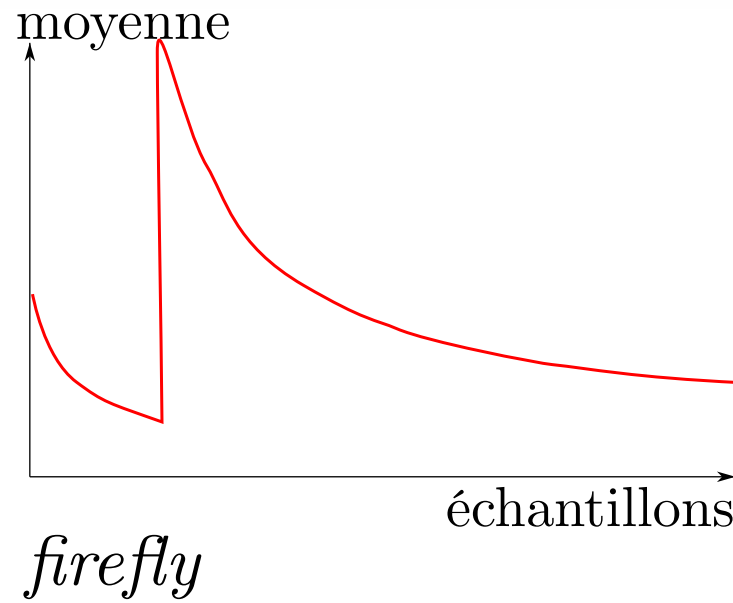
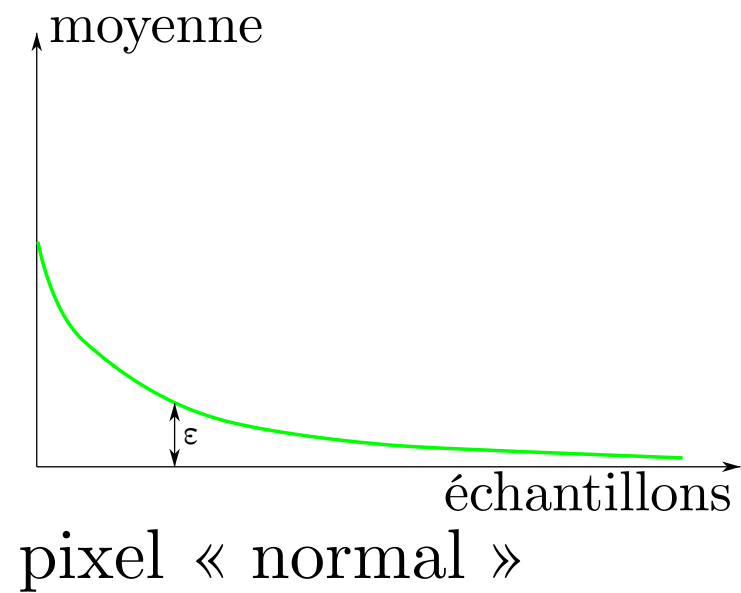
- Buisine et al. (2021) ICMLA

Comparaison des modèles

(à paraître Springer, 2022)



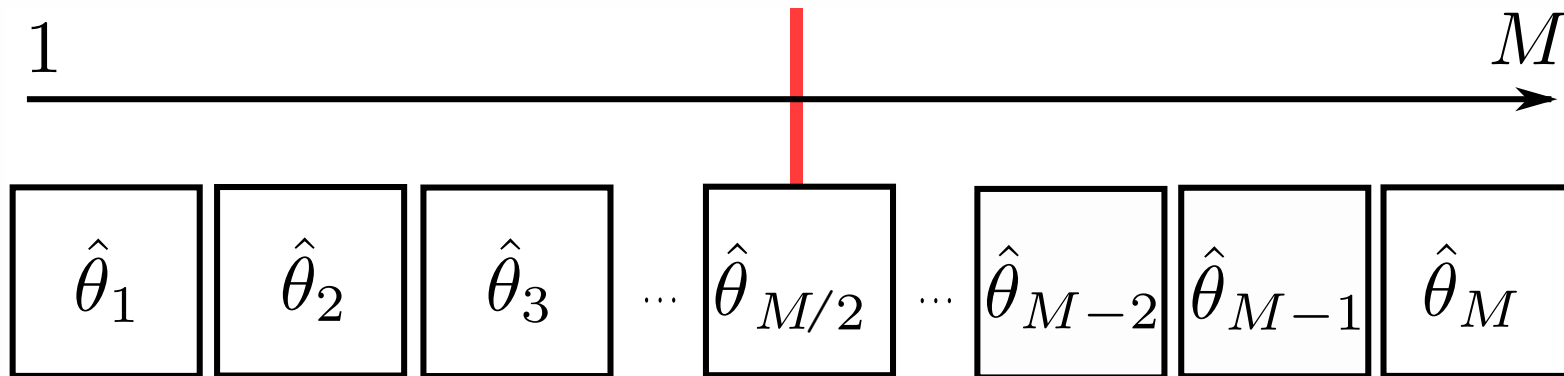
Un cas extrême du bruit : le firefly



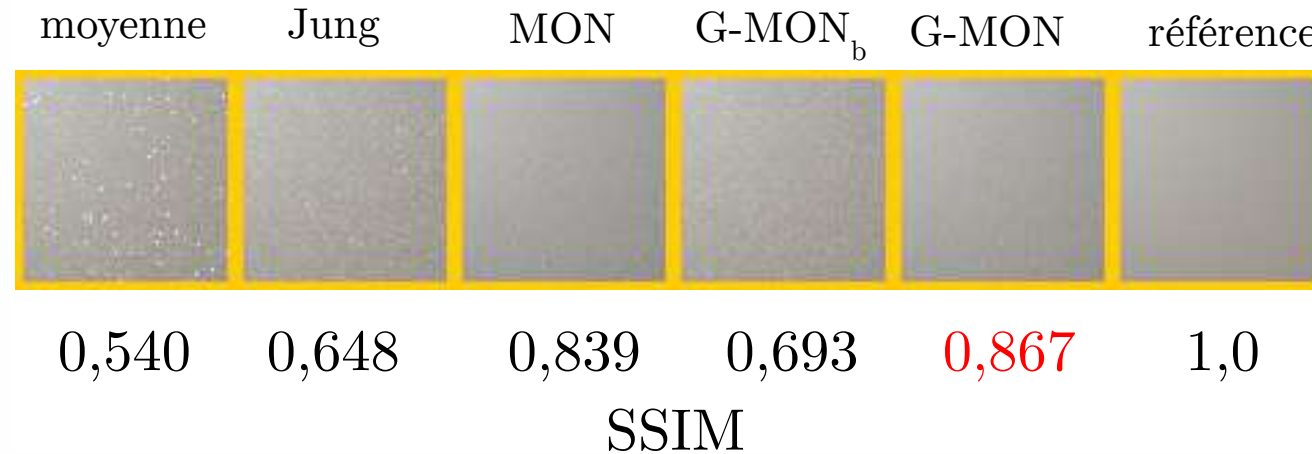
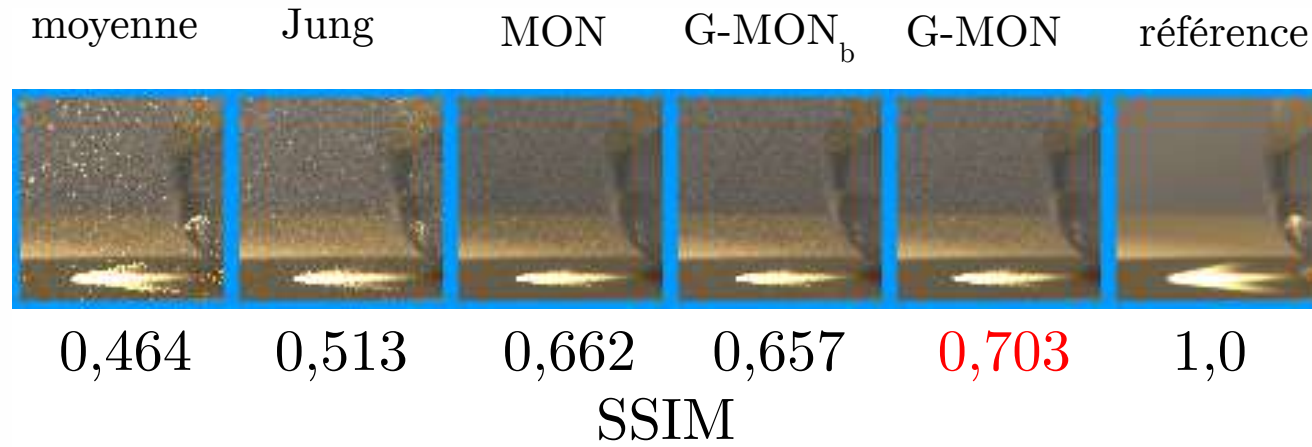
Solutions basées sur le MON

Constat

- moyenne : convergence rapide, facile à calculer mais peu robuste
- médiane : robuste mais coûteux et convergence plus lente (tri, stockage)
- médiane des moyennes (MoN): estimateur intéressant



MoN et G-MoN : résultats



Buisine, Delepouille & Renaud (2021) EGSR

Plan

1. Problématique générale
2. Rendu photoréaliste
3. Nos travaux
4. **Perspective / Conclusion**

Prolonger les travaux entamés

- affiner les comparaisons
- meilleurs estimateurs
- généraliser à d'autres moteurs
- zones (meilleure segmentation de l'image)

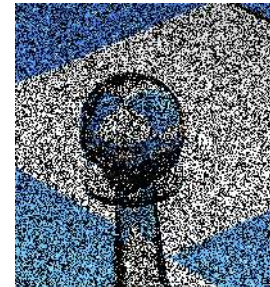
Les nouveaux enjeux de la perception numérique

Débruitage

- Approche complémentaire
- Étapes de filtrage + information des pixels voisins + estimations
- Intel® OpenImageDenoiser



1 spp



10 000 spp



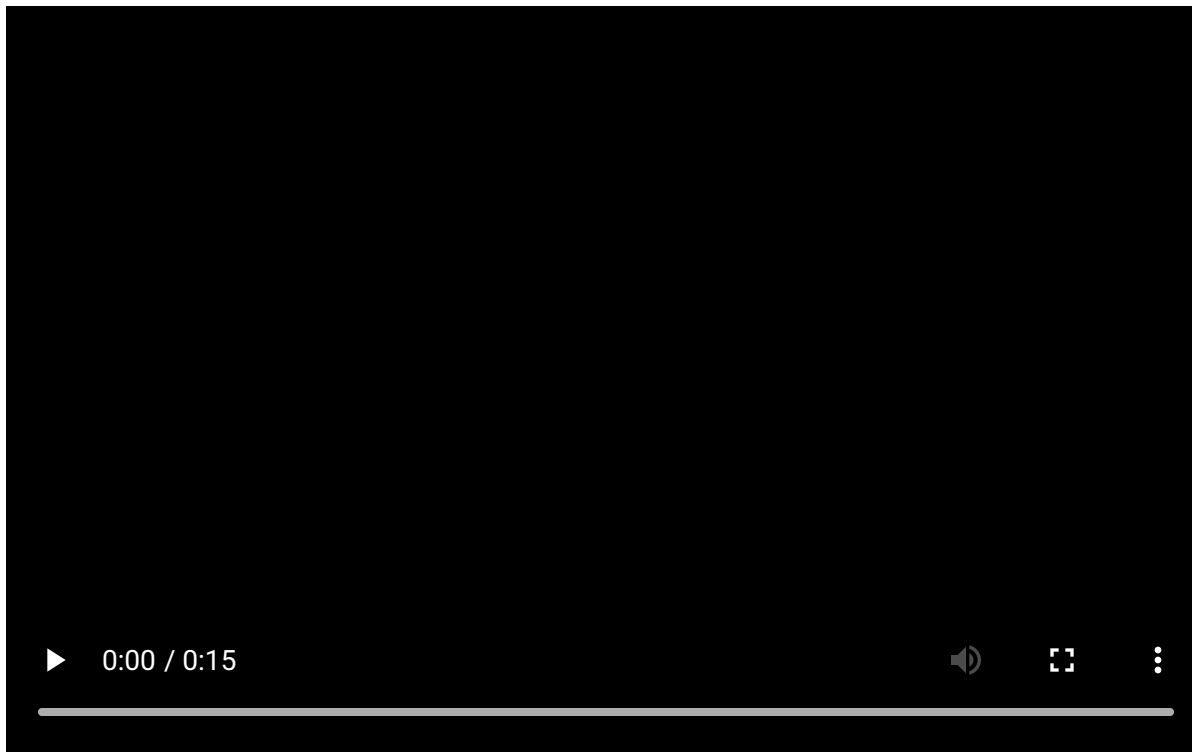
sans débruitage



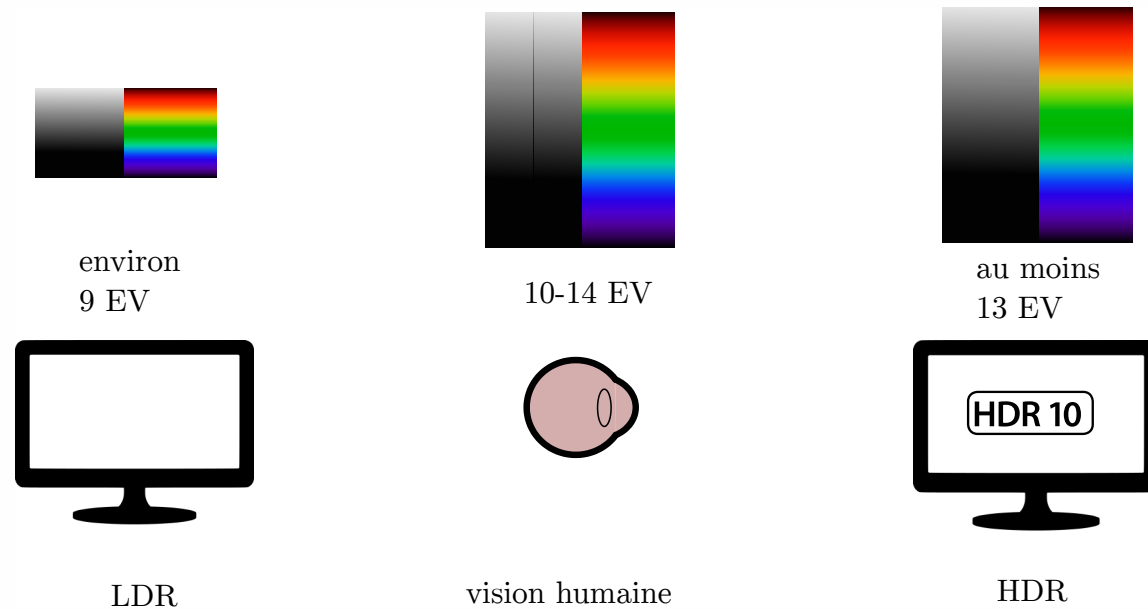
avec débruitage

Séquences d'images : vidéo

- facteur d'échelle : 1 min. de vidéo = 1500 frames
- problèmes de cohérence temporelle
- très forte sensibilité temporelle du SVH



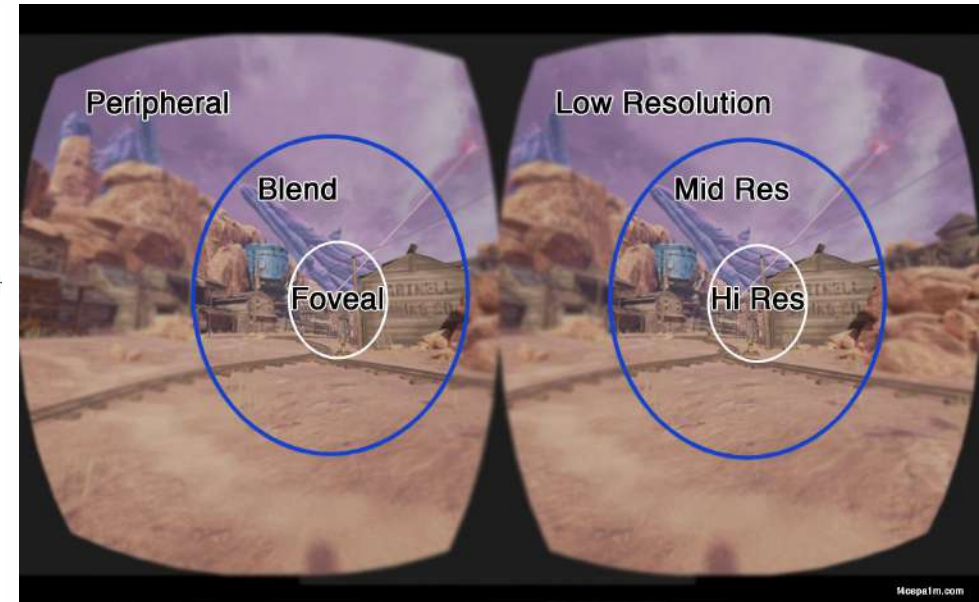
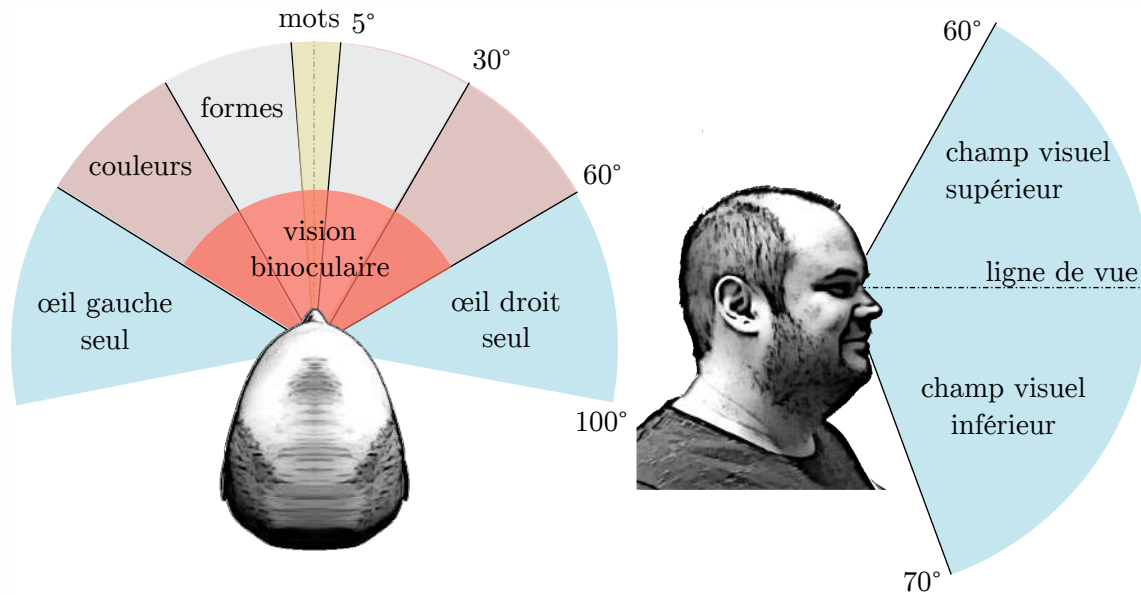
images HDR = High Dynamic Range



- luminance = facteur clé de la perception du bruit

Rendu fovéal

- problème de la résolution des casques RV
- principe : ne calculer précisément que dans la zone d'acuité maximale
- calculs moins précis dans les zones périphériques



- conditions
 - zone de transition
 - bon suivi de la zone fovéale

Conclusion

- Utiliser les données perceptives pour toutes les étapes des calculs
- Ébauche d'une étude intégrée de la perception dans un cadre d'images numériques
- Approche trans-disciplinaire
- Tout ceci n'est qu'un début ...

A detailed Renaissance-style landscape painting. The scene is divided into several distinct areas. In the foreground, a rocky cliffside is the site of a battle. A large, reddish-brown fire burns on the right, with a figure falling from a high rock. Below, soldiers on horseback and on foot are engaged in combat. In the middle ground, a coastal town with a harbor is visible. The harbor is filled with various ships, including galleons and smaller boats. The town features a prominent church with a tall spire and several castles or fortified buildings. In the background, a vast landscape with rolling hills and distant mountains is visible under a pale sky. The overall style is characteristic of the Northern Renaissance, with fine detail and a rich color palette.

Début de la question