

# Evaluation en Visualisation d'Information

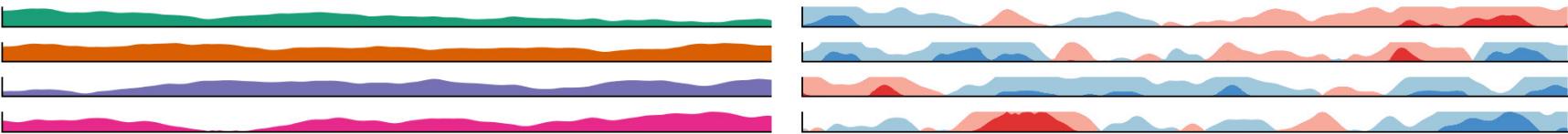
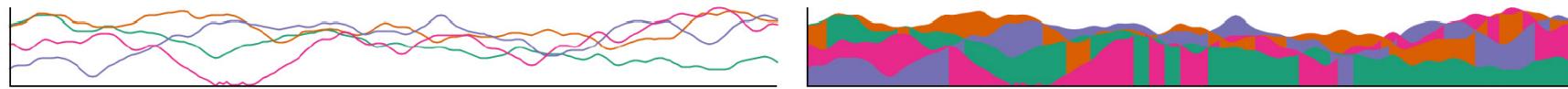
G.-P. Bonneau

# Evaluation en Visualisation d'Informations

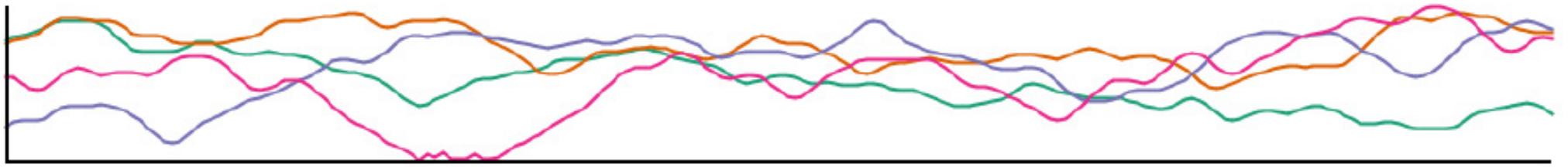
- Evaluation par tâches
- Méthodes psychophysiques

# Evaluations par tâches

- Evaluation comparée de 4 méthodes de visualisation de séries temporelles multiples

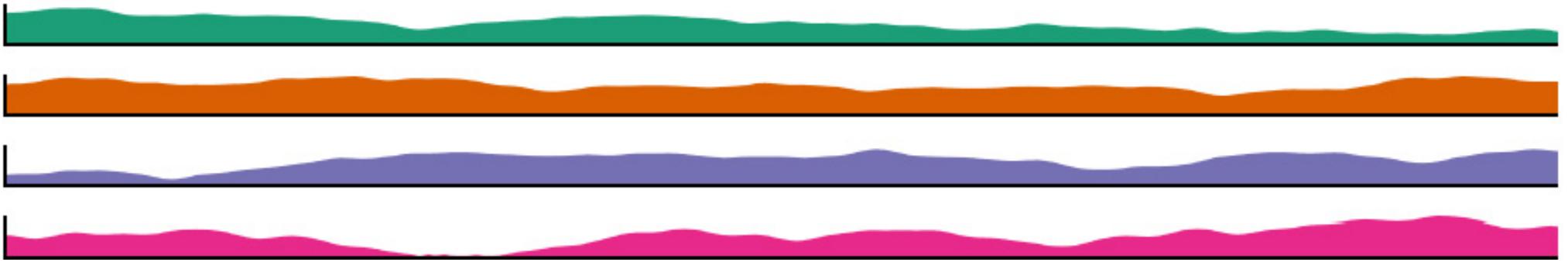


# Graphes simples superposés



L'espace disponible est partagé

# Small multiples



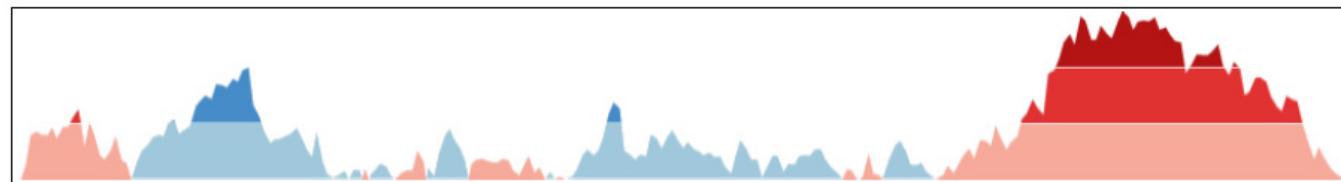
L'espace disponible est divisé

# Graphe Horizon

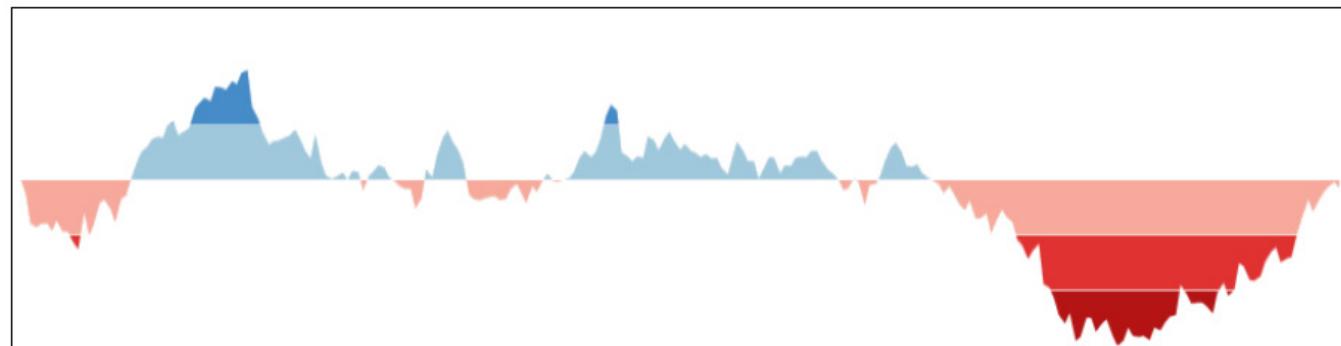
Ecrasement



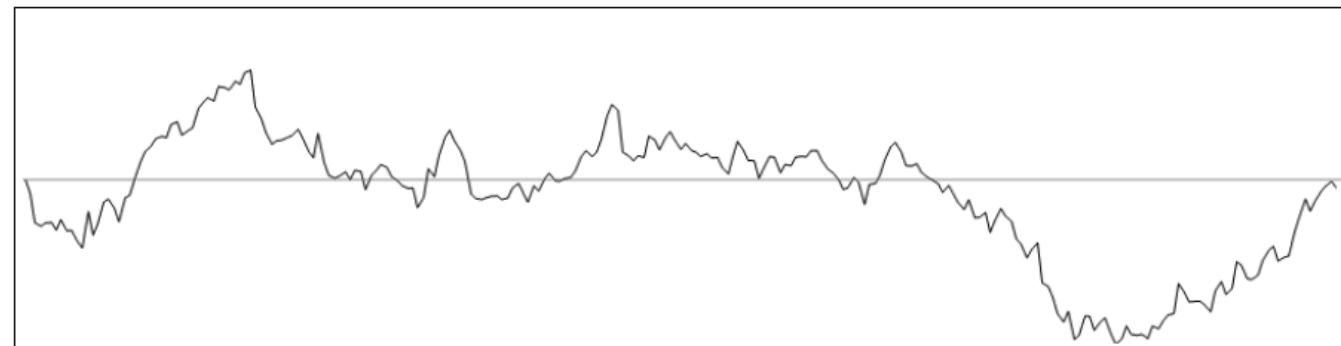
Miroir



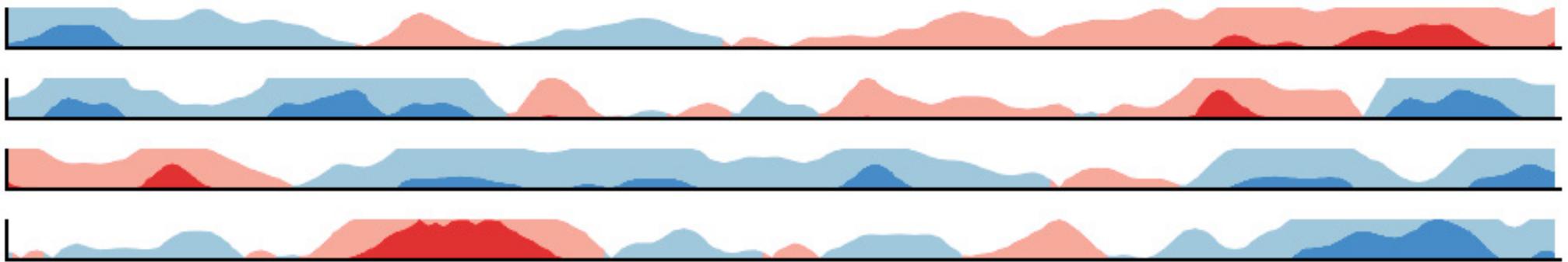
Découpage coloré



Données



# Multiples graphes horizons



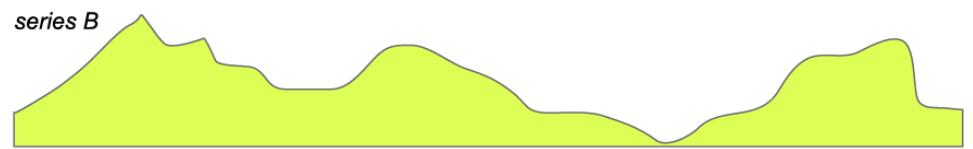
# Graphes tressés (Braided Graph)



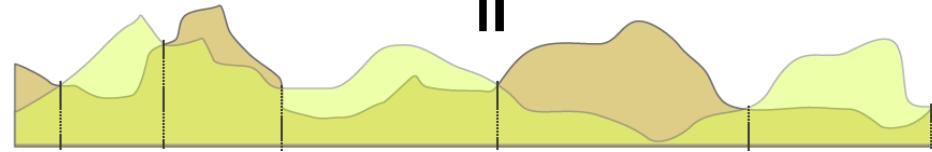
# Graphes tressés



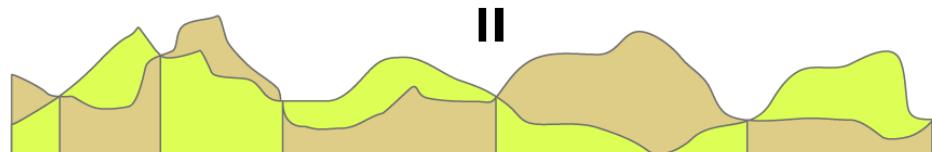
+



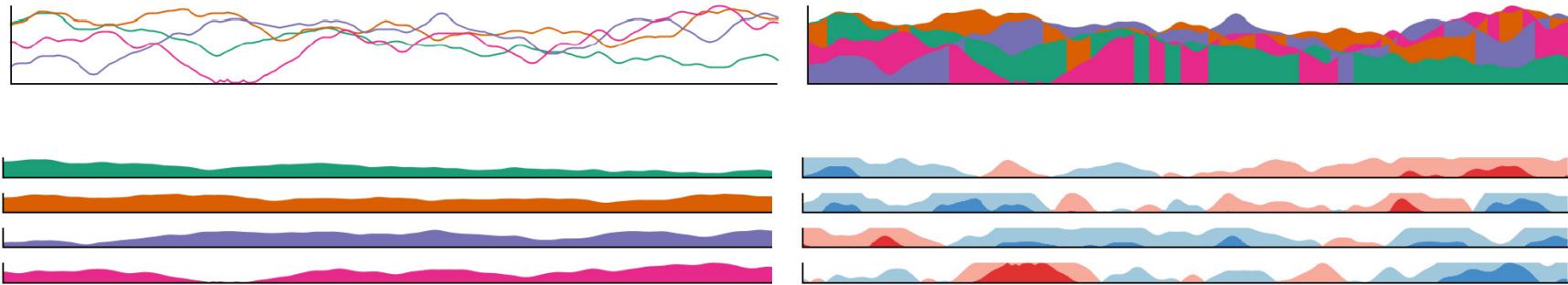
Segmentation  
aux intersections



Dessin dans l'ordre des  
décroissant des tailles



# Hypothèses



- H1: les méthodes en espace partagés donnent de meilleurs résultats pour des tâches locales.
- H2: les méthodes en espace divisés donnent de meilleur résultats pour des tâches globales.
- H3 : les performances décroissent quand le nombre de séries augmente.
- H4: L'espace vertical disponible influence directement les résultats

# Tâches

- Max:
  - Trouver la valeur maximum en un même point
    - Tâche locale (localisée horizontalement)
- Slope:
  - Trouver la série avec la plus forte croissance
    - Tâche globale (nécessite de visualiser sur toute la largeur)
- Disc:
  - Trouver la valeur maximum en des points différents
    - Tâche globale

# Expérimentation

- Description des participants (nombre, âge, connaissances informatiques, correction visuelle...)
- Description du matériel utilisé, des couleurs, des labels
- Visualisations statiques, réponses données par une simple lettre désignant une série
- Séries générées aléatoirement pour chaque test
- Une seule visualisation par écran

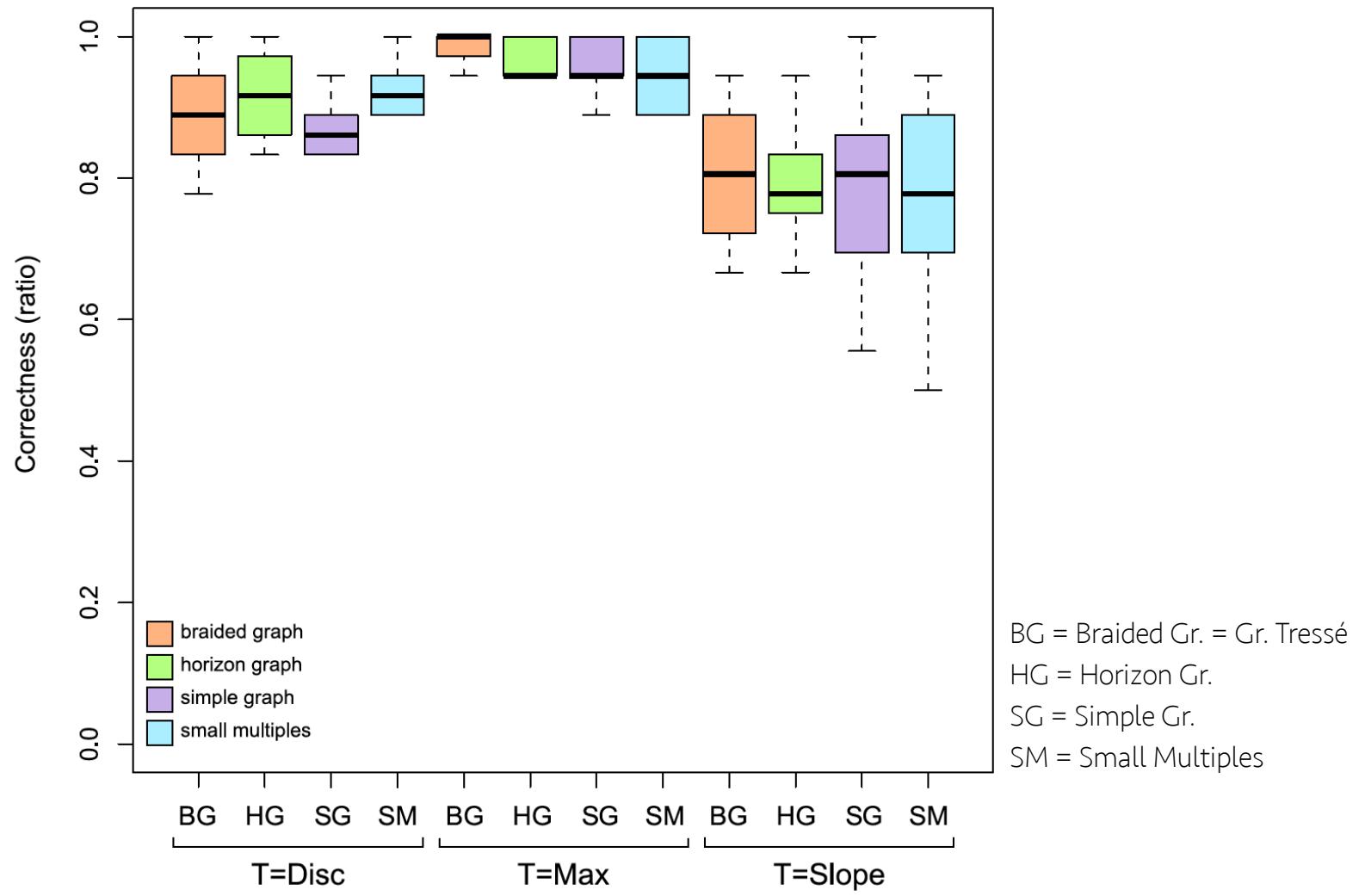
# Nombre de cas possibles

- 4 méthodes de visualisation (=V)
- 3 nombres différents de séries (2, 4 et 8) (=N)
- 3 tailles d'affichage (=S)
- 3 tâches à réaliser (=T)

$$V * N * S * T = 108 \text{ cas possibles}$$

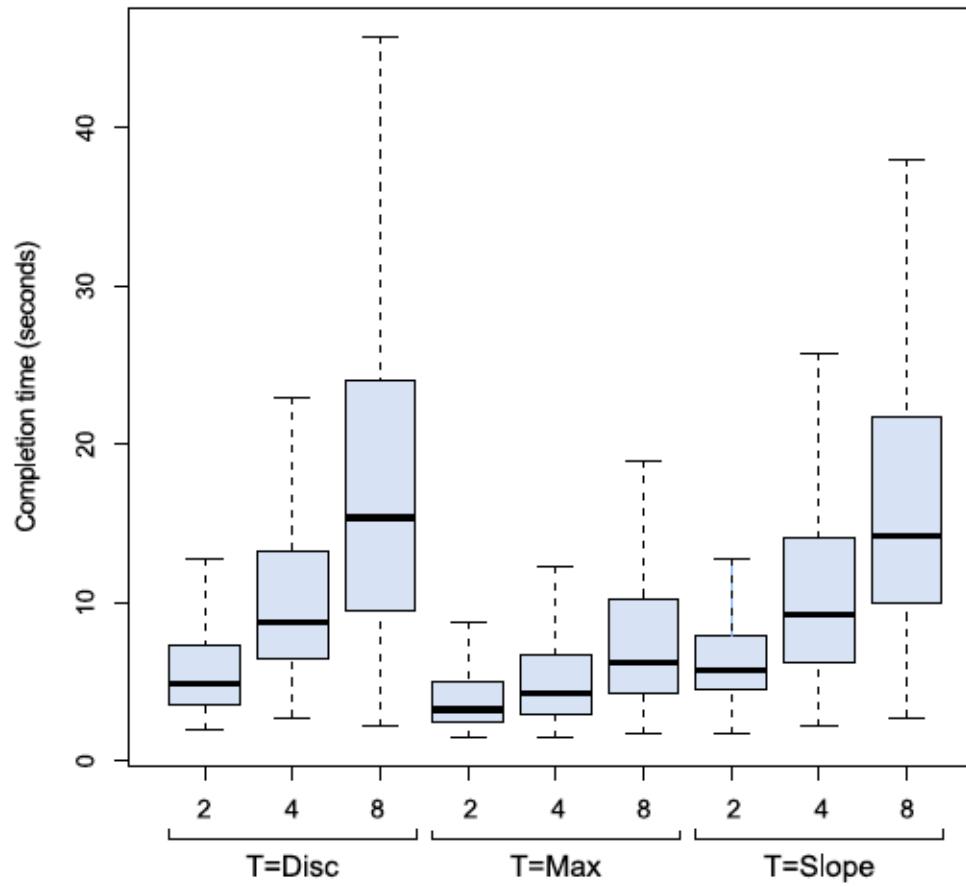
- 16 participants réalisant 2 fois chaque cas:  
 $16 * 2 * 108 = 3456$  tests pour l'expérimentation

# Résultats



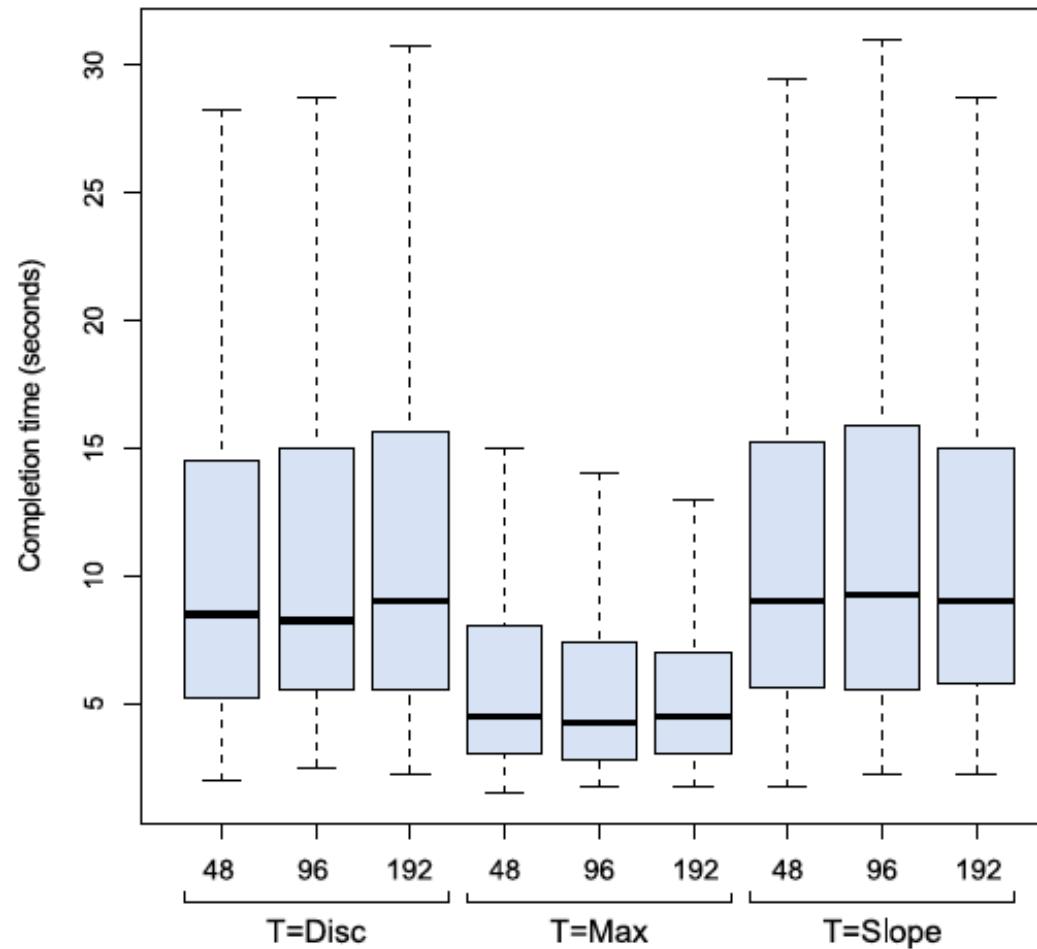
- Les taux de bonne réponse sont statistiquement semblables

# Résultats



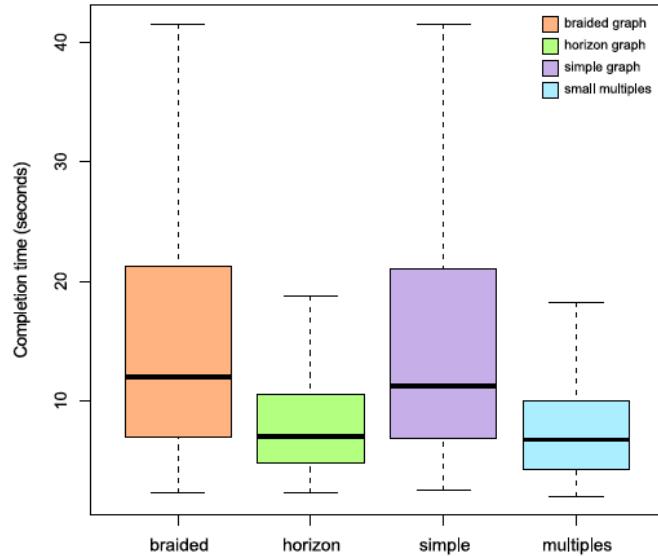
- Le nombre de séries influence fortement le temps de réponse (H3)

# Résultats

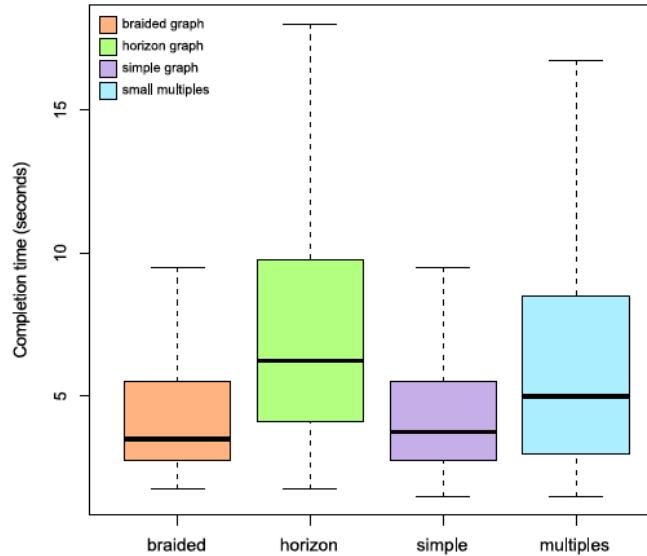


- La taille d'affichage n'influence pas le temps de réponse (contredit H4)

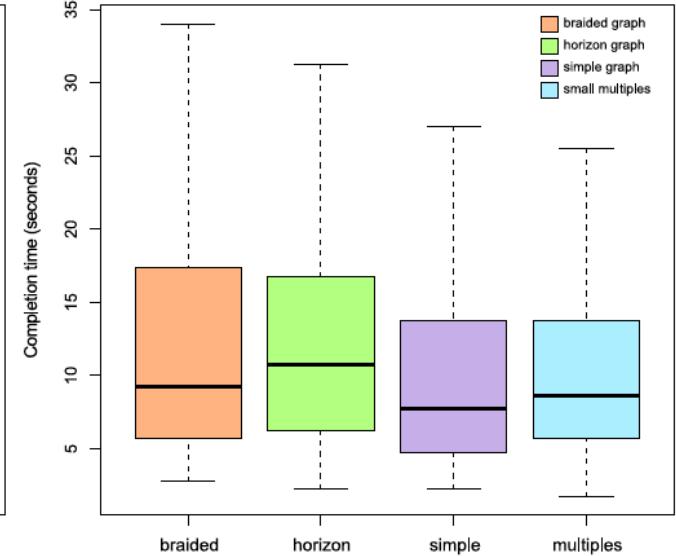
# Résultats



Tâche Disc (globale)



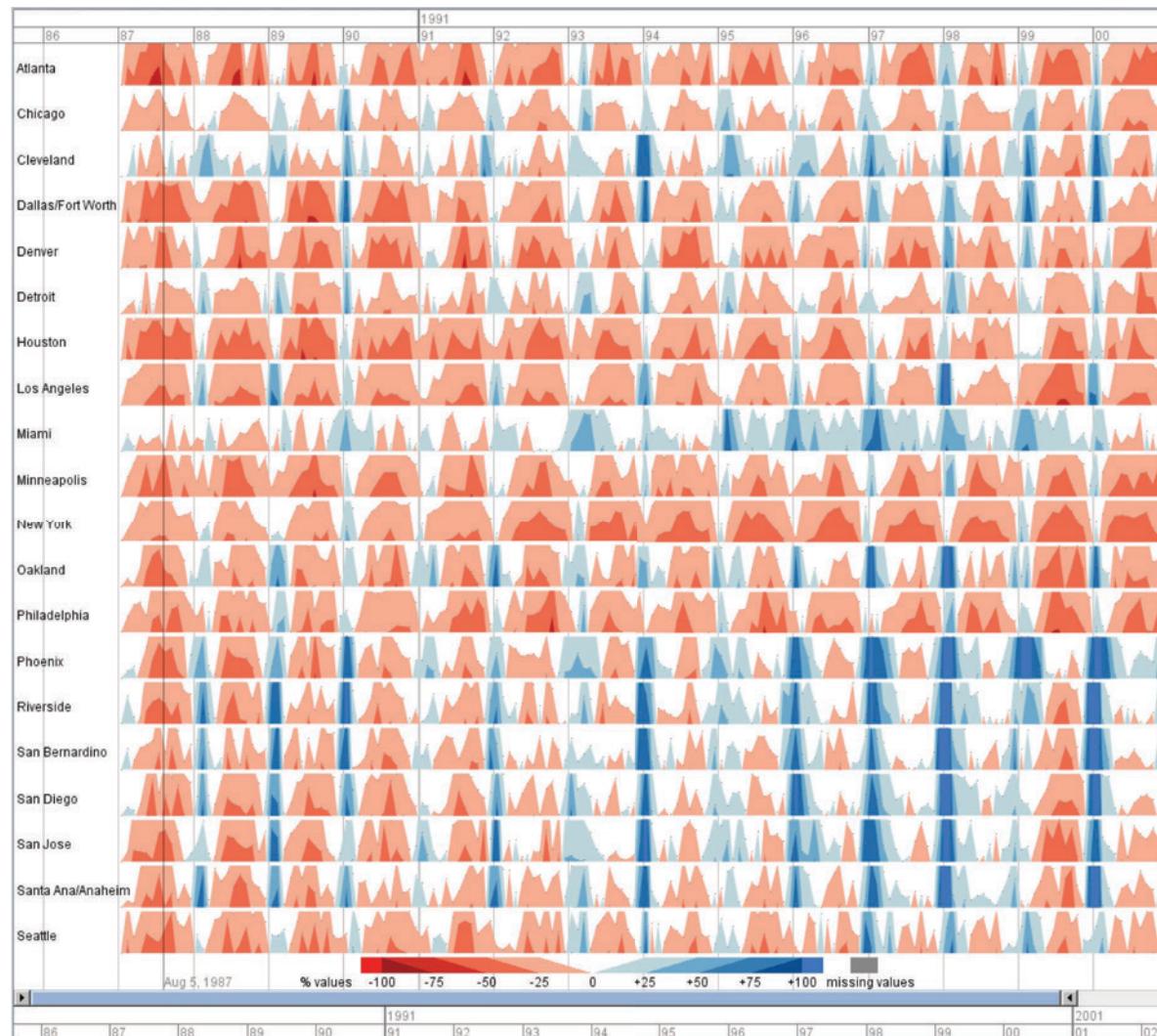
Tâche Max (locale)



Tâche Slope (globale)

- Les méthodes en espace partagés (simple, braided) sont plus rapides pour la tâche locale (H1)
- Les méthodes en espace divisés (horizon, multiples) sont plus rapides uniquement pour la tâche globale Disc (H2)

# Généralisation à des grands nombres de séries



- Les conclusions précédentes sont elles encore valables?

# Méthodes Psychophysiques

Méthodes physiques de mesures de la perception sensorielle

Adaptées à l'étude des limites du système visuel

Nécessite une définition physique rigoureuse du stimuli visuel

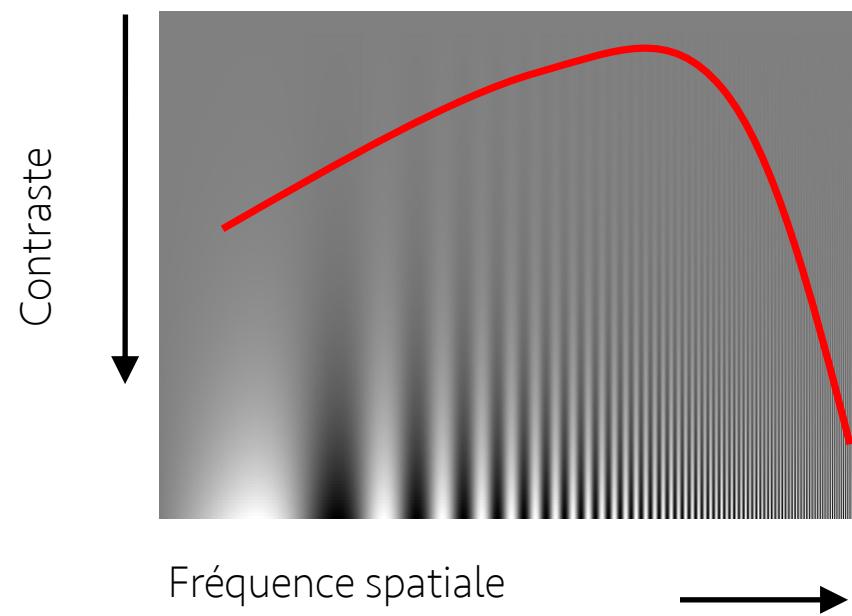
Difficile d'en déduire des règles pratiques

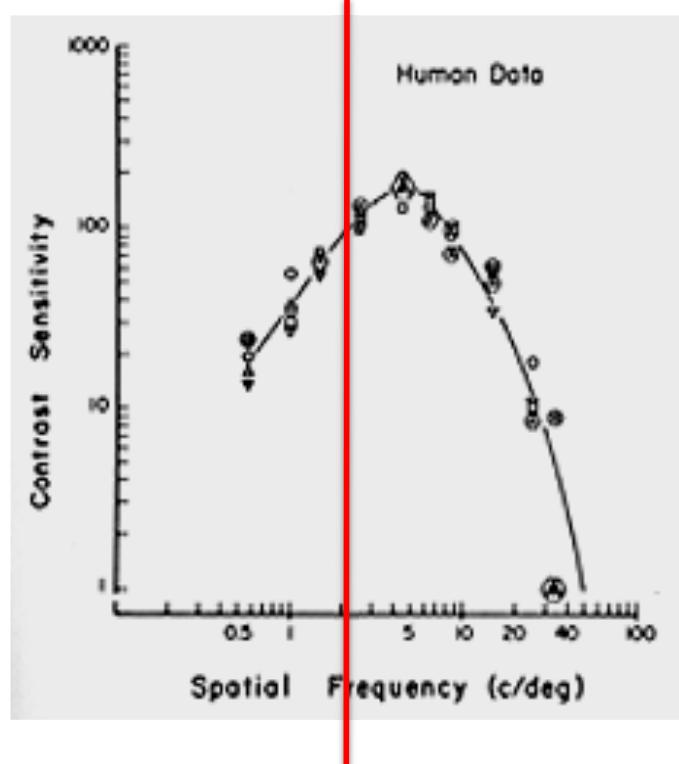
# Méthodes Psychophysiques

Exemple:

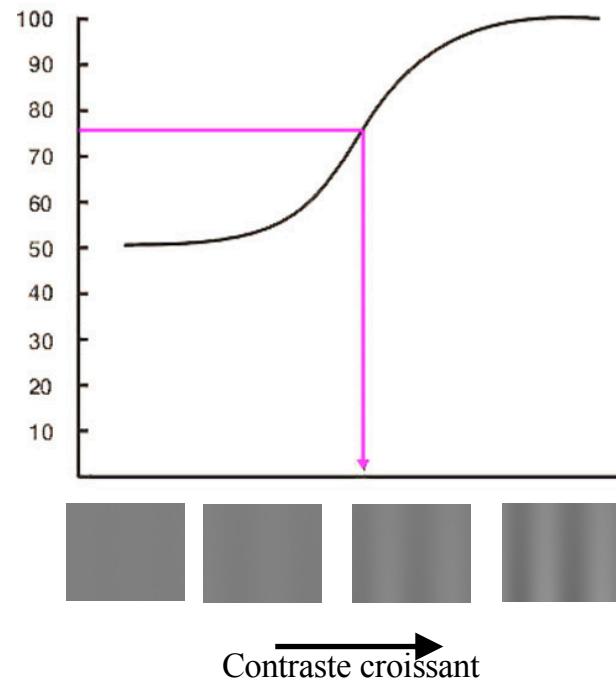
mesure des seuils de sensibilité au contraste

# Fonction de Sensibilité au Contraste



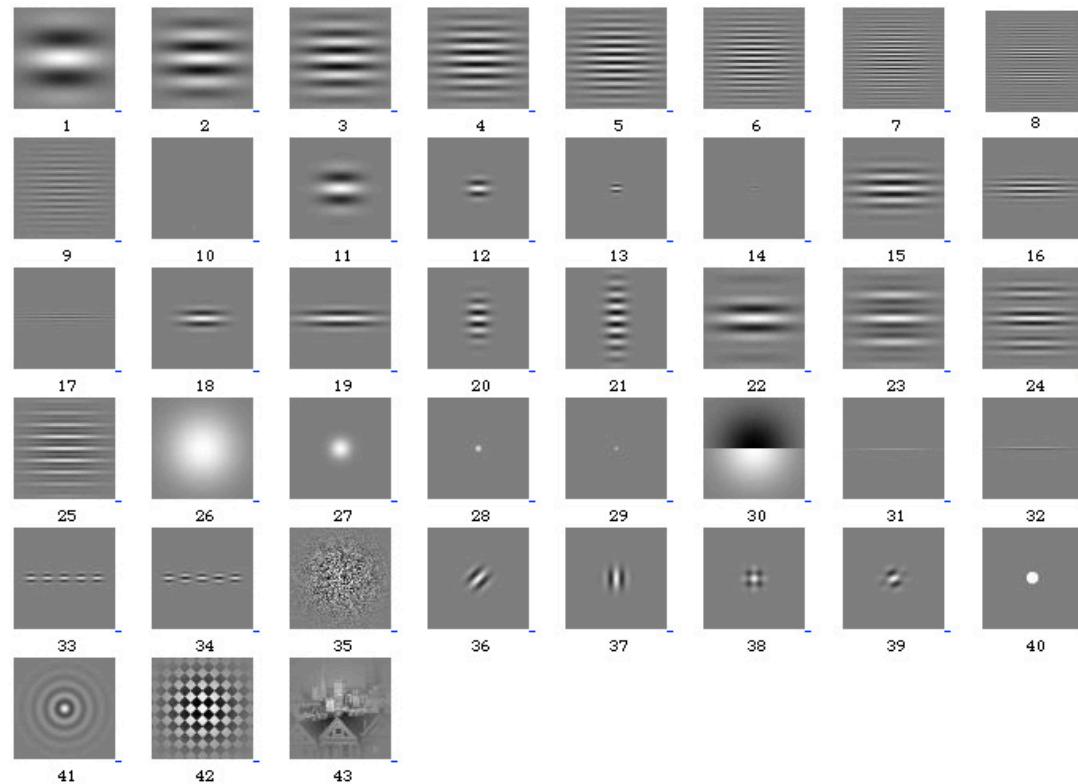


DeValois & DeValois (1988)  
*Spatial Vision*. Oxford: Oxford University Press



# The ModelFest data base

- <http://vision.arc.nasa.gov/modelfest/>



Valeurs de seuil pour  
44 stimuli  
16 observateurs  
12 laboratoires

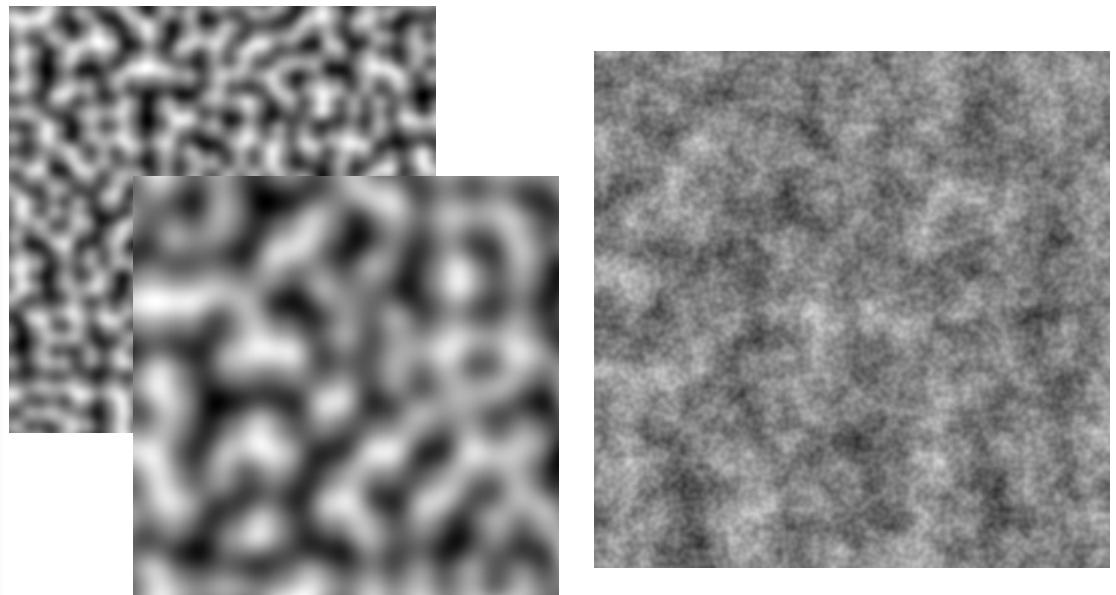
Méthode  
2IFC  
256x256 pixel  
stimuli  
120 pixels par  
degré  
0.5s  
présentation

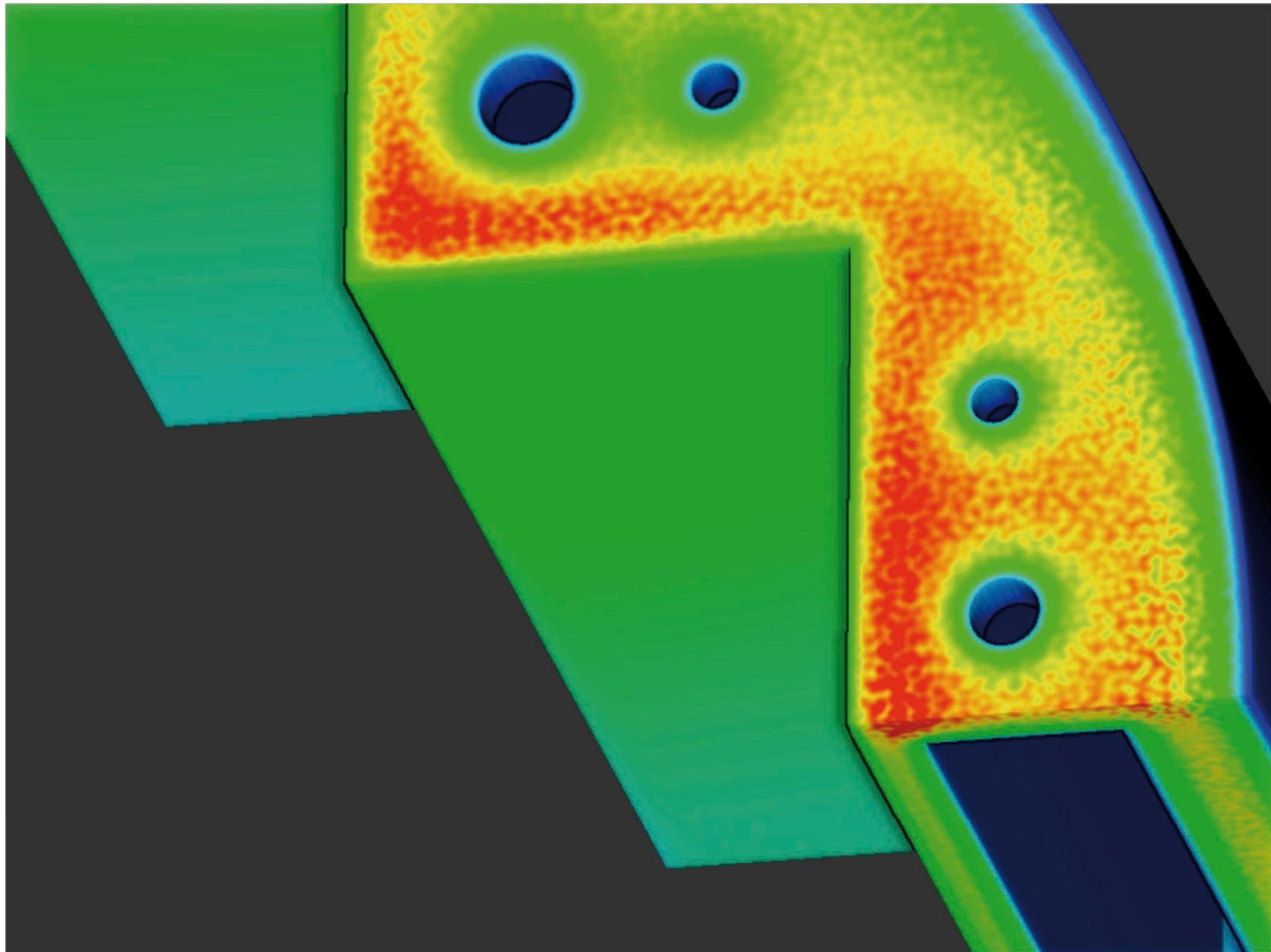
...

[detail](#)

# Bruit de Perlin

$$n_{f_0,p}(\vec{x}, t) = \sum_i p^i N(f_0 2^i \vec{x}, t)$$





Coninx, Bonneau & al., APGV 2011

# Expérimentation psychophysique

## Seuil de sensibilité au contraste pour le bruit de Perlin

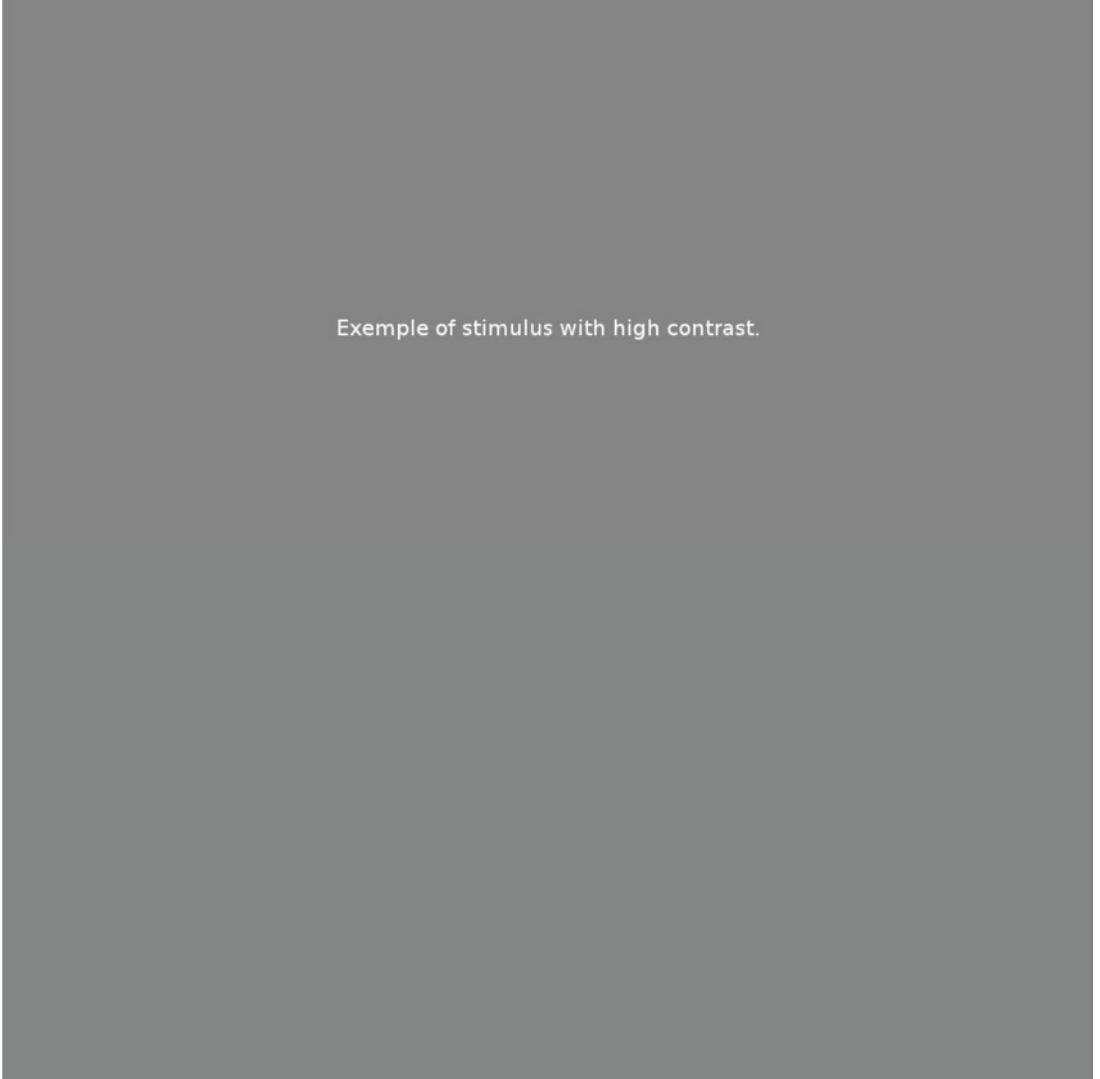
- Choix de 4 valeurs de fréquence  $f_0$  et de persistence  $p$   
=> 16 stimuli
- $f_0 = \{ 2, 4, 8, 16 \}$     $p = \{ 0., 0.25, 0.5, 0.75 \}$
- Calcul de stimuli de taille 256x256
- LCD avec pixel de taille 0.295mm, Distance écran-sujet  
2m
  - => 120 pixels par degré
- 30 essais pour la mesure d'un seuil
  - =>  $30 * 16 = 480$  essais, ~5s chaque essai => ~40mn de durée d'expérimentation par sujet

# Pièges des expérimentations psychophysiques

- Nécessité de calibrer les réponses en luminance de l'écran
- Le delta de niveau de gris doit être  $< 0.25\%$  du seuil recherché pour une détection fiable
- ...
- Il est préférable de collaborer avec des experts

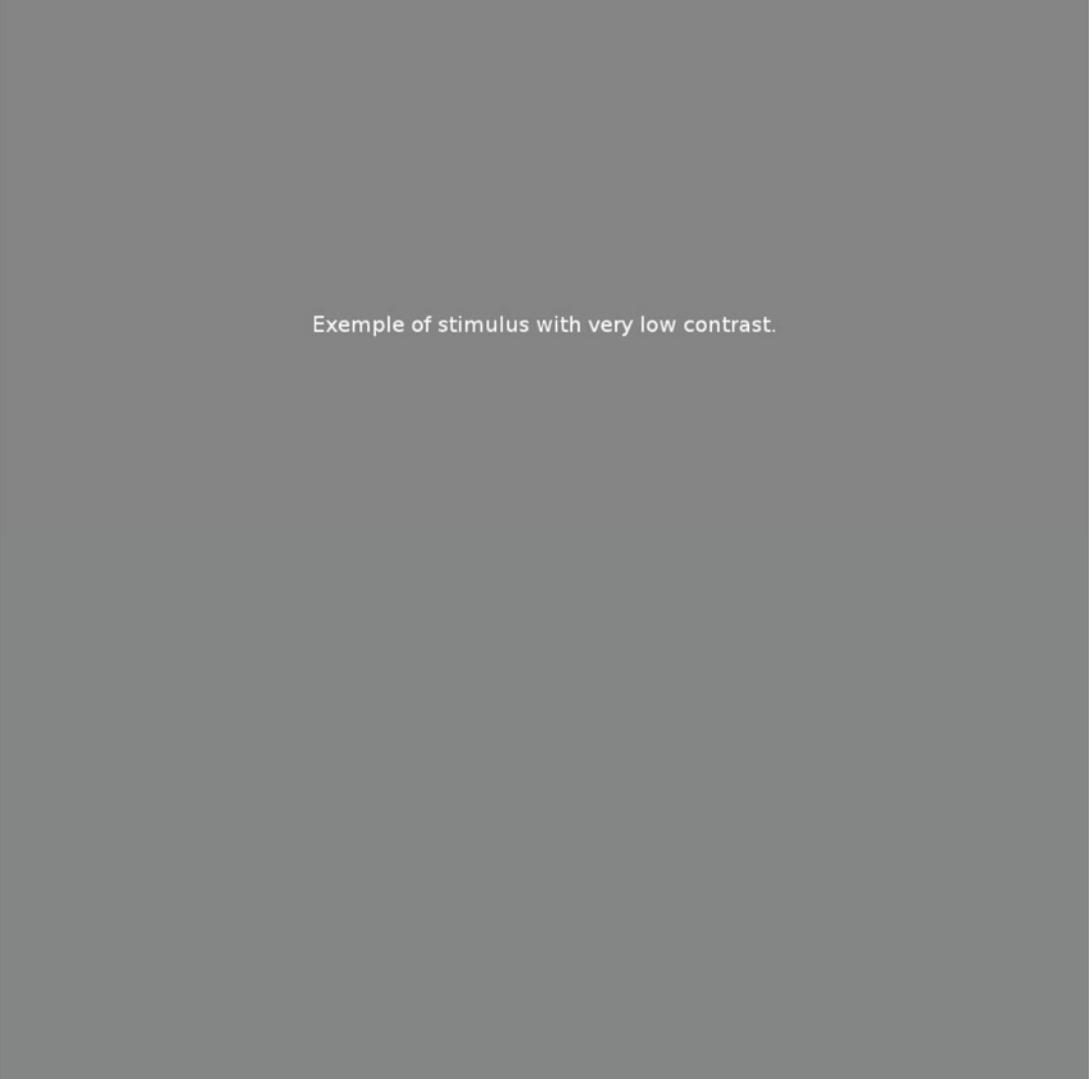
# Stimulus de fort contraste

Exemple of stimulus with high contrast.



# Stimulus de faible contraste

Exemple of stimulus with very low contrast.

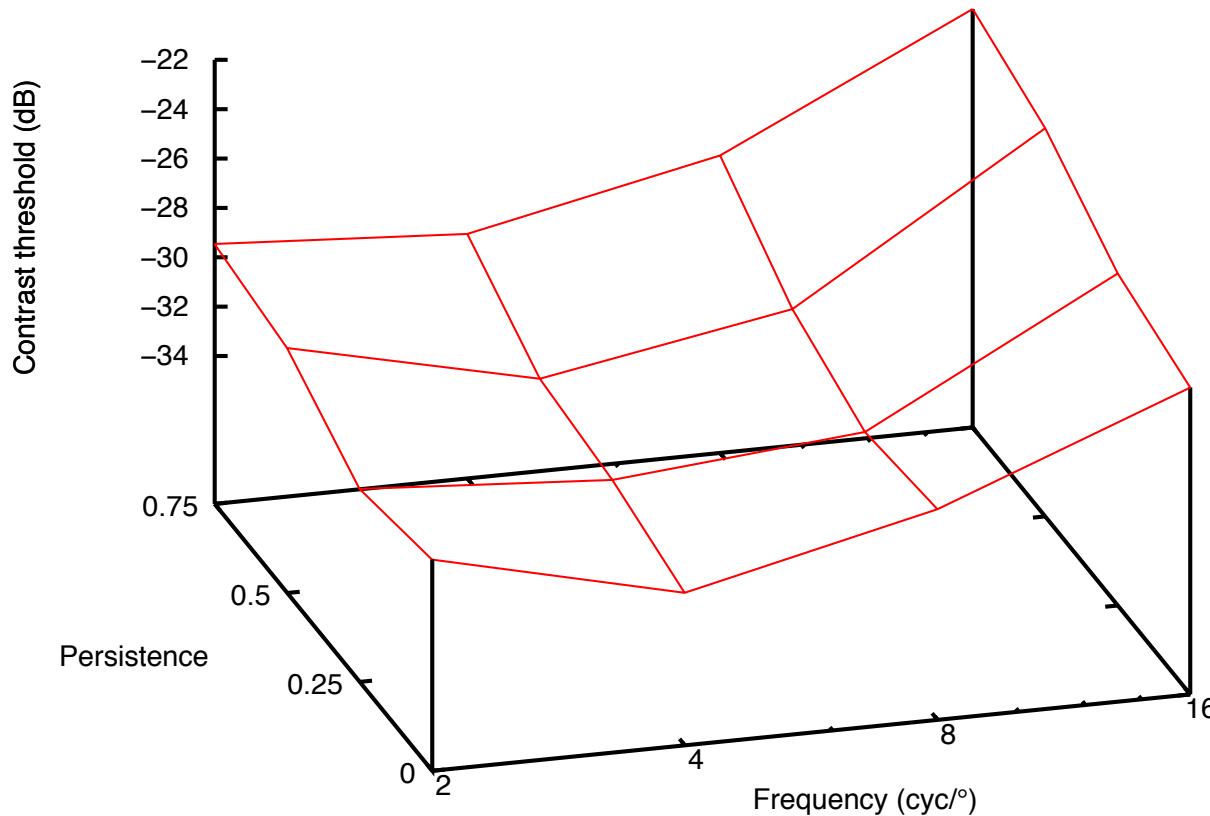


# Contraste proche du seuil

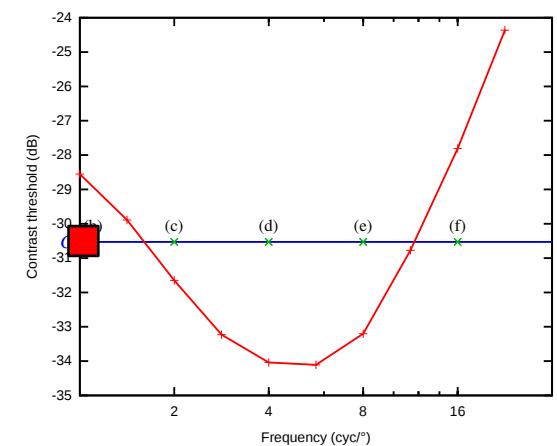
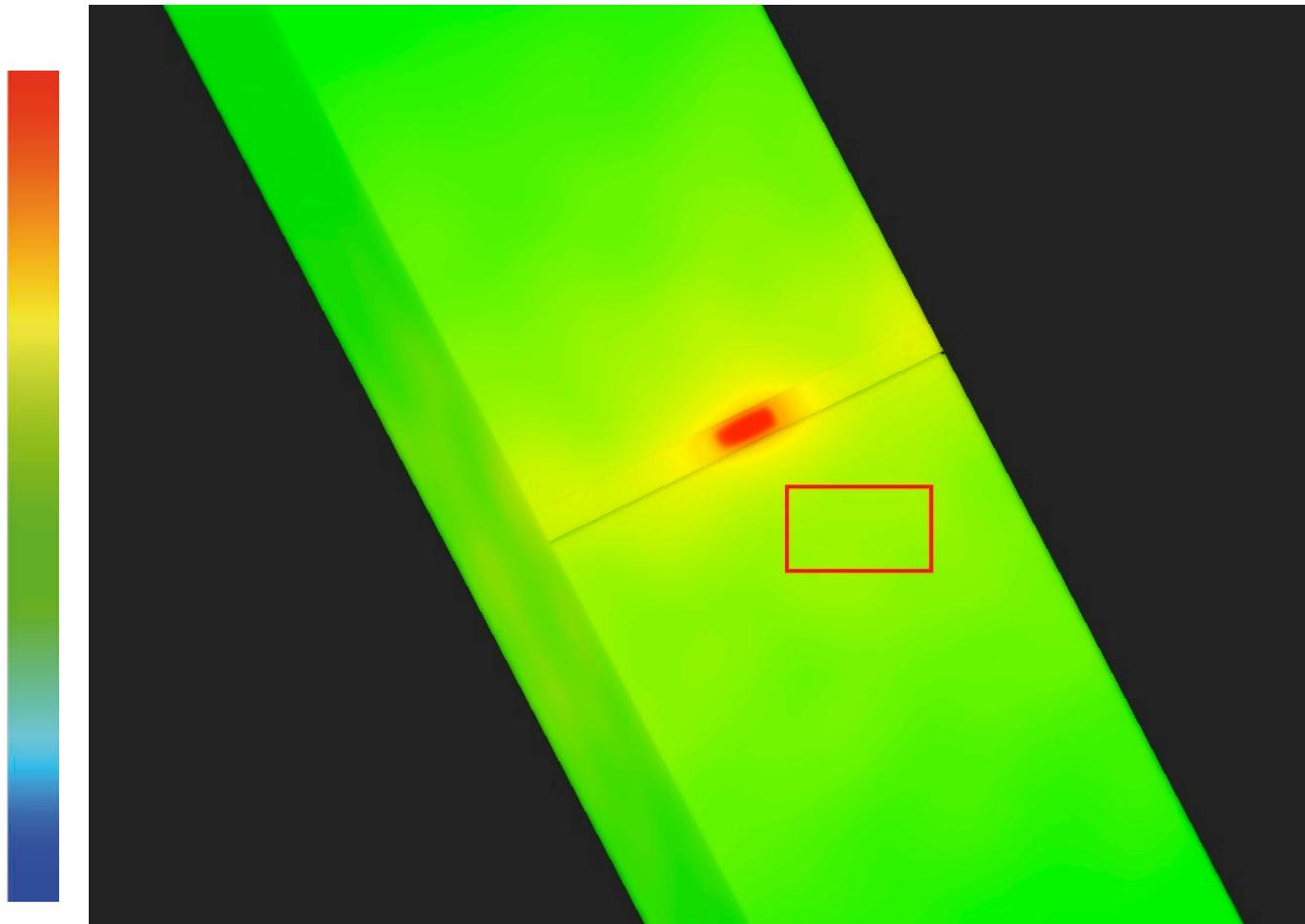
Exemple of stimulus with near-threshold contrast.



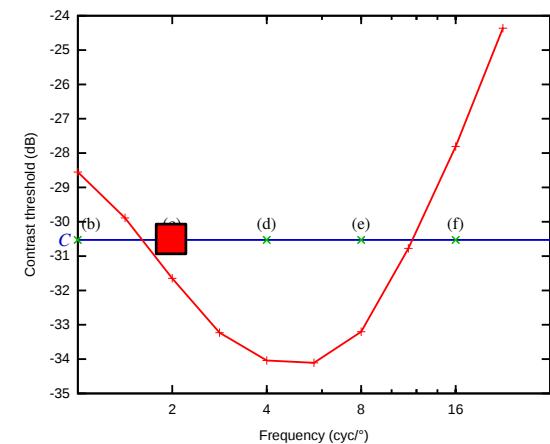
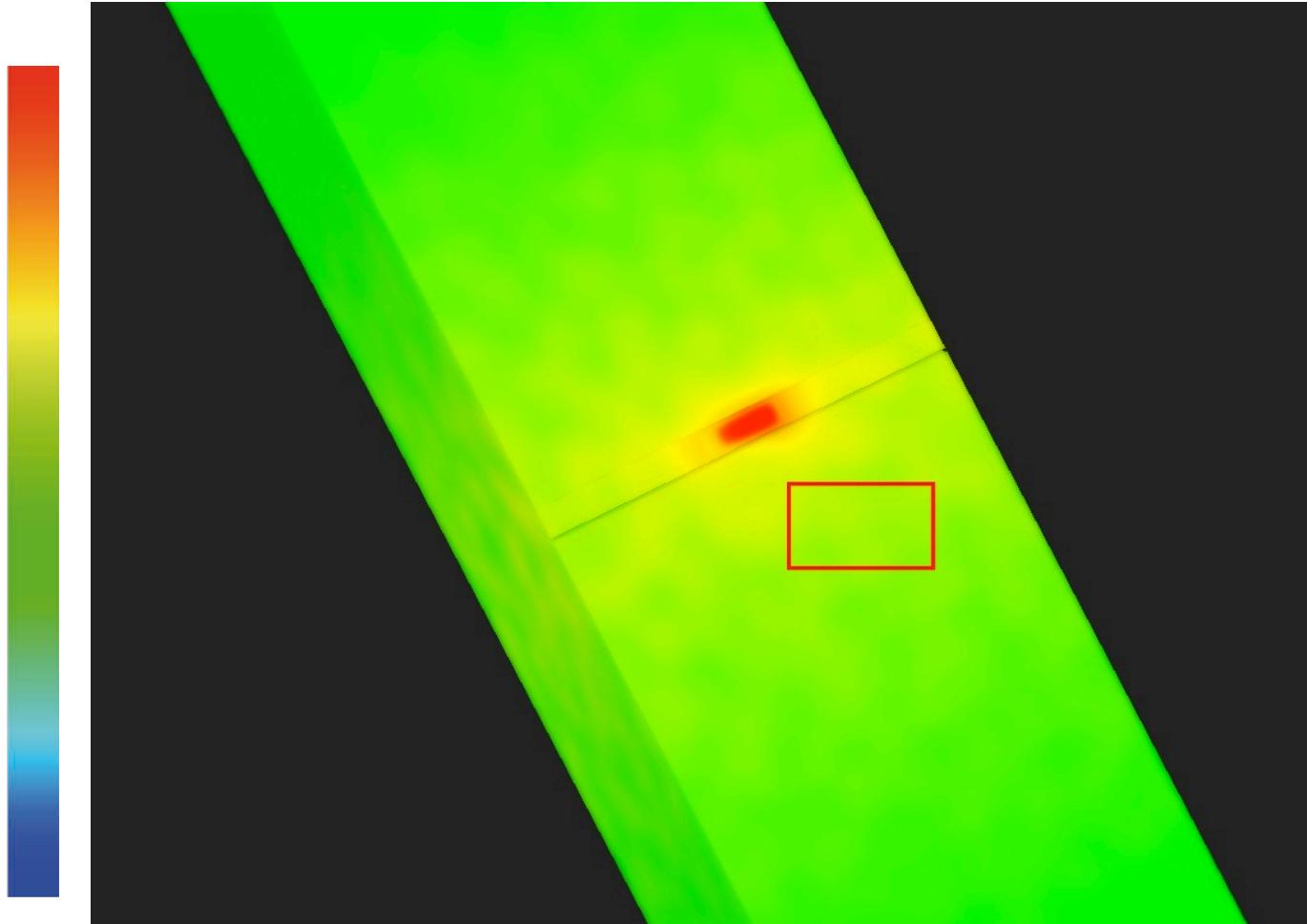
# Résultat de l'expérimentation Psychométrique



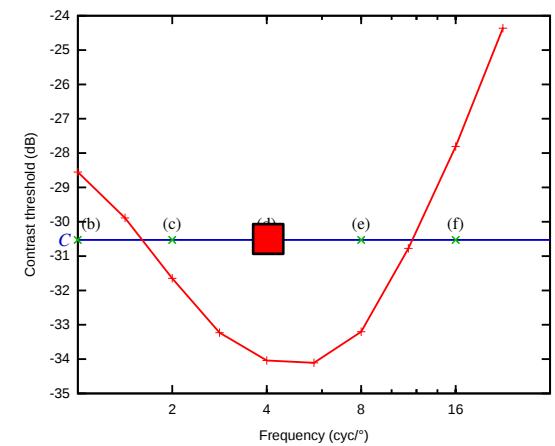
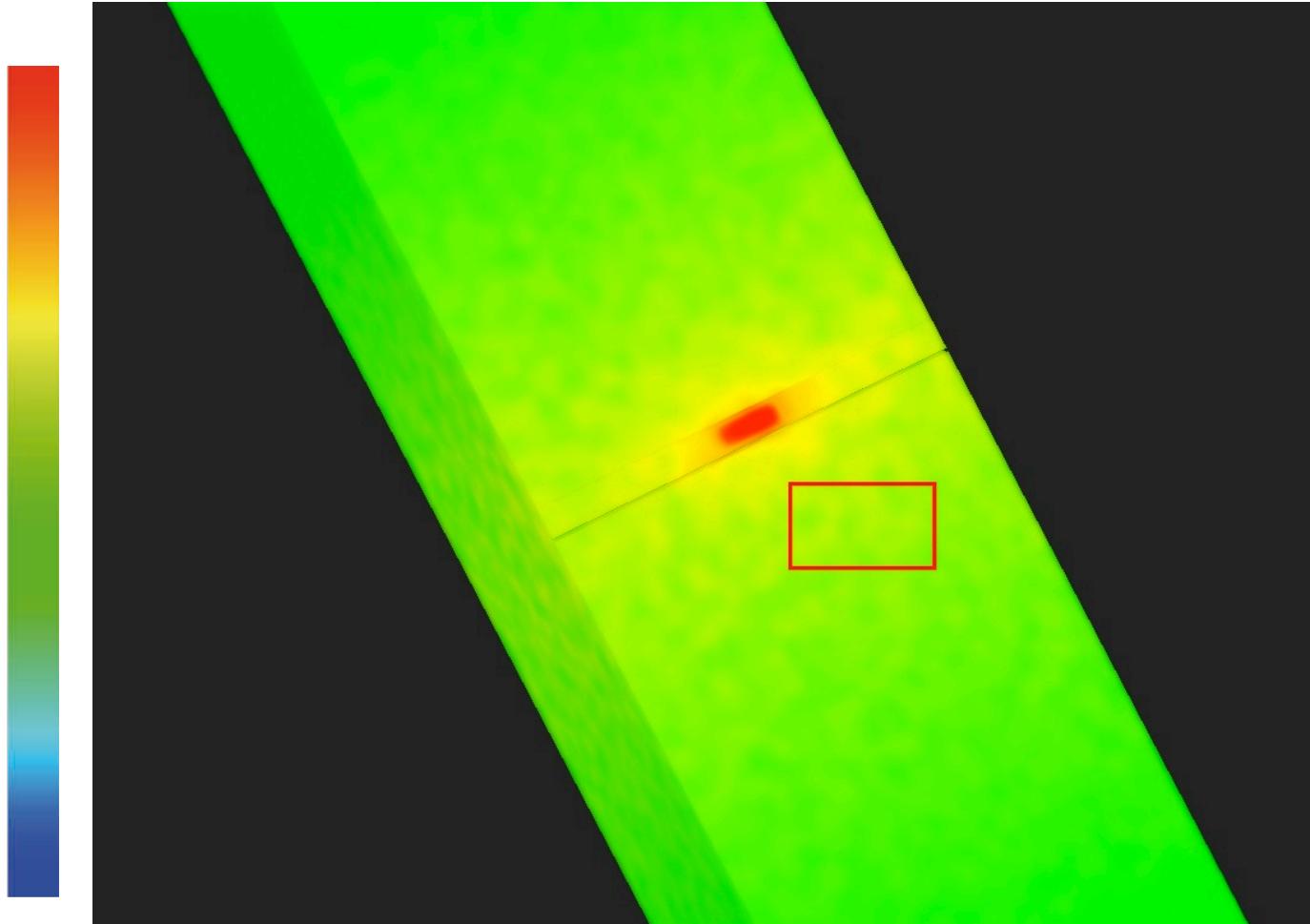
# Utilisation des résultats expérimentaux



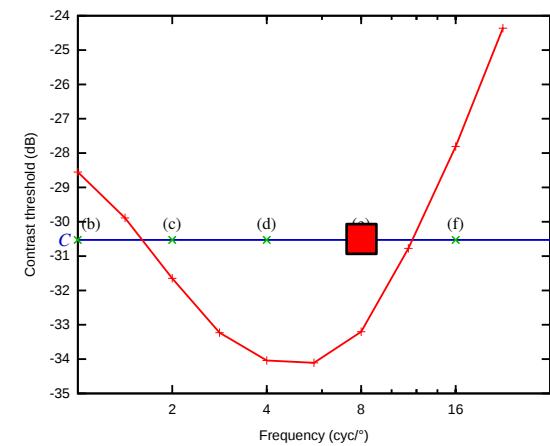
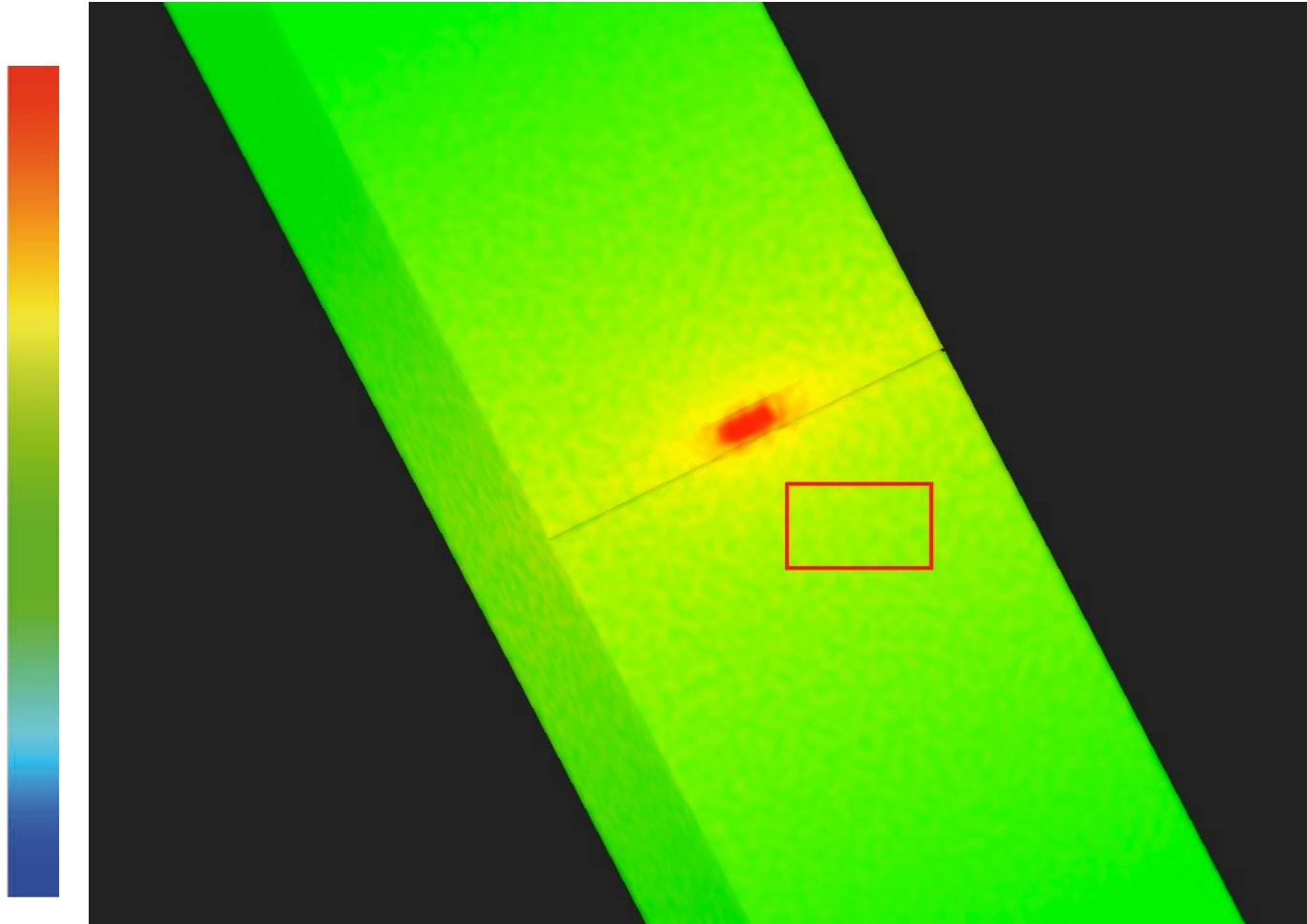
# Utilisation des résultats expérimentaux



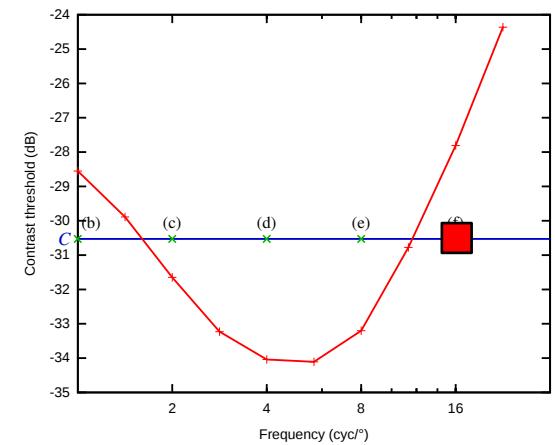
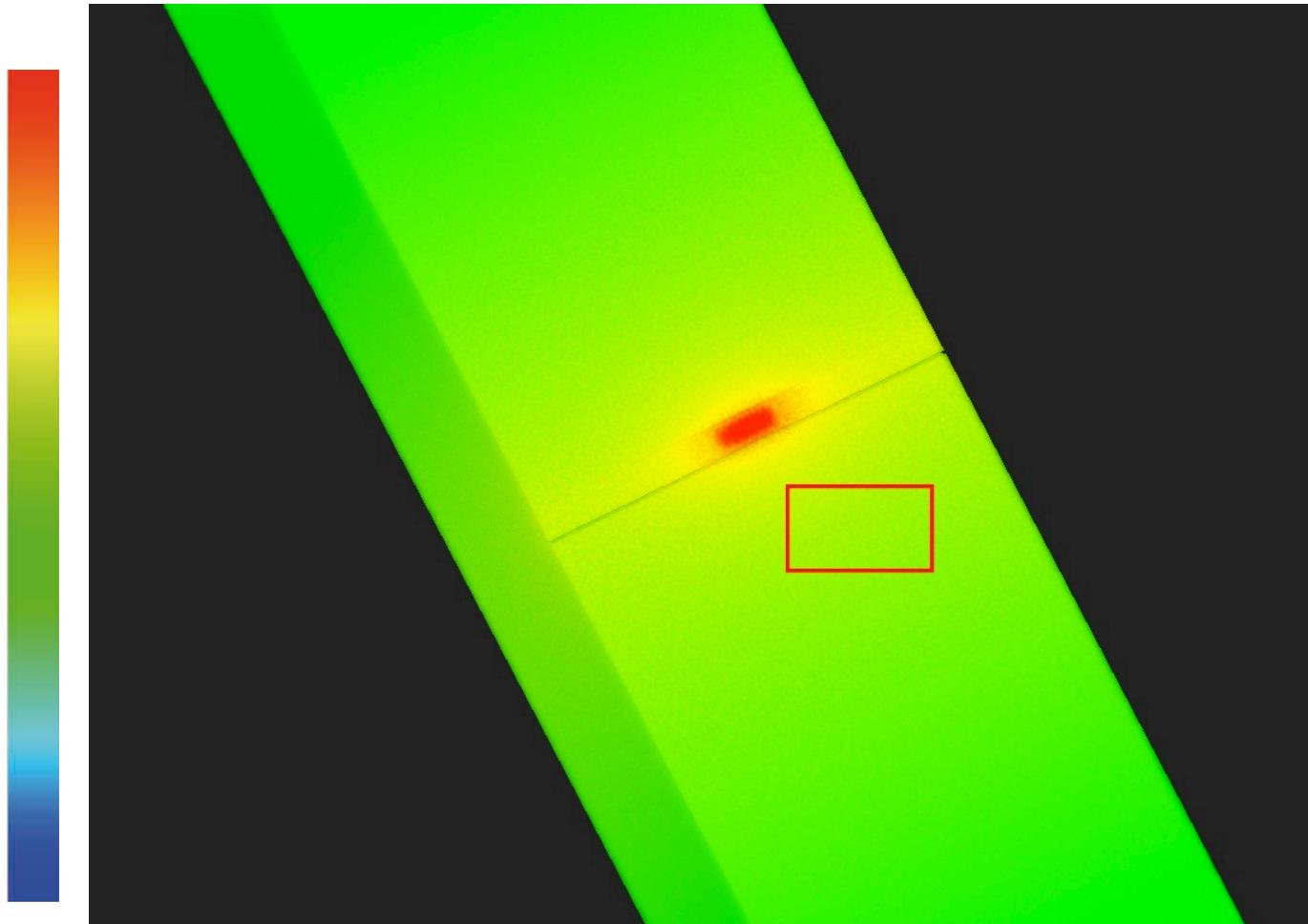
# Utilisation des résultats expérimentaux



# Utilisation des résultats expérimentaux



# Utilisation des résultats expérimentaux



# Attention!

- Expérimentation niveaux de gris, utilisation couleur
- Expérimentation stimuli avec environnement neutre, utilisation dans environnement complexe

# Conclusion

- Evaluation par tâches cognitives
- Evaluation par expérimentation psychométriques
- L'évaluation est nécessaire, elle est objective
- Elle produit des critères de choix d'une visualisation adaptée à une application