

RESPECT DES PERFORMANCES VISUELLES LORS DE LA VISUALISATION D'IMAGES DE SYNTHÈSE

R. Brémond, E. Dumont
Laboratoire Central des Ponts et Chaussées, Paris

Résumé

Le LCPC développe des techniques de synthèse d'images permettant de réaliser des études de visibilité et de lisibilité routière, notamment sur simulateur de conduite. Le principal obstacle identifié à l'utilisation de telles images porte sur les procédés de visualisation sur écran des images de synthèse calculées, du fait des limites techniques des systèmes de restitution. Cet obstacle est contourné, en pratique, par des algorithmes de transposition permettant de convertir un signal visuel à afficher en un signal affichable par le système de restitution utilisé. Les algorithmes utilisés sont des algorithmes de transposition de luminance (*tone mapping*). Il résulte de l'utilisation d'un tel algorithme une dégradation du signal, et des stratégies sont proposées dans la littérature pour minimiser cette dégradation. Nous présentons une méthodologie qui a été élaborée et mise en œuvre dans le but d'évaluer la qualité de ces algorithmes en termes de performance visuelle des observateurs, pour des niveaux lumineux correspondant à une situation de conduite diurne.

1 Introduction

Le LCPC développe des techniques de synthèse d'images permettant de réaliser des études de visibilité et de lisibilité routière, notamment sur simulateur de conduite. Ces images sont définies en unités photométriques et colorimétriques (exemple Fig. 1), permettant ainsi de décrire quantitativement le signal visuel perçu par un observateur. Actuellement, le principal obstacle identifié à l'utilisation de telles images porte sur les procédés de visualisation sur écran.



Figure 1 - Carte d'iso-luminances calculée à partir d'un modèle des effets visuels du brouillard (Dumont, 2002).

Nous avons cherché à évaluer la capacité des procédés de visualisation existants à respecter les performances visuelles des observateurs, ce qui est le critère visuel le plus pertinent en matière de visibilité routière, à l'étude de laquelle ces images sont destinées. L'objectif est donc de quantifier l'écart entre les images présentées et les scènes modélisées, de manière à élargir le domaine de validité des images de synthèse pour les études de visibilité et de lisibilité routière.

2 Limites des systèmes de visualisation d'images

2.1 Calibrage des écrans

Le premier problème à résoudre lors de l'affichage sur écran d'une image en unités physiques est de s'assurer que les valeurs physiques des couleurs et des luminances affichées correspondent aux valeurs prévues. Pour cela, il est nécessaire de caractériser précisément le système de restitution. Pour un écran CRT¹, par exemple, l'affichage des images se base sur des valeurs d'adressage (des triplets d'entiers) correspondant aux photophores *Rouge*, *Vert* et *Bleu* de l'écran : c'est le triplet {RVB}. Pour contrôler les images affichées, il est nécessaire de connaître la correspondance entre un triplet quelconque {RVB} et le triplet {XYZ} qui sera observé sur l'écran. En pratique, cette étape repose sur une caractérisation photométrique et colorimétrique de l'écran utilisé. Pour un écran CRT, on utilise classiquement le modèle GOG (*gain-offset-gamma*) préconisé par la CIE (CIE, 1996). Il consiste à modéliser la luminance affichée par un moniteur CRT comme la somme des luminances des trois canaux, en décrivant la luminance d'un canal par une fonction puissance de la valeur d'adressage (modulée par un gain) et par un bruit de fond (*offset*). Cette formulation est inversible, ce qui permet de calculer la valeur d'adressage nécessaire pour afficher à l'écran une couleur XYZ donnée. Cette procédure permet de soumettre l'observateur au même stimulus visuel que ce qui était prévu dans l'image exprimée en unités physiques.

2.2 Transposition de luminances

Les systèmes de restitution présentent des limites techniques incontournables, qui font que l'affichage des valeurs en unités physiques (calculées ou mesurées) n'est pas toujours possible. Ces limites sont de plusieurs ordres (limitation du domaine de couleur, effets de la quantification des valeurs d'adressage, stabilité dans le temps, etc.). La principale limitation, avec les technologies actuelles, est que la luminance disponible est souvent beaucoup plus faible que celle qui est présente dans la scène à afficher. De la même manière, la dynamique de luminance² des écrans est souvent trop faible de plusieurs ordres de grandeur par rapport aux besoins. Ces limitations sont contournées, en pratique, par des algorithmes de « transposition de luminance » (*tone mapping*), permettant de convertir un signal visuel à afficher (de luminance L) en un signal visuel différent L' , affichable par le système de restitution utilisé. Il résulte globalement de l'utilisation d'un algorithme de transposition une dégradation du signal, et des stratégies sont proposées dans la littérature pour minimiser cette dégradation (Tumblin *et al*, 1993 ; Ward, 1994 ; Larson *et al*, 1997).

3 Évaluation de la qualité d'une transposition de luminance

La dégradation des images calculées due aux limites des systèmes de restitution est susceptible d'avoir pour conséquence une modification des performances visuelles des observateurs, par comparaison avec les performances qu'ils auraient avec un système de restitution « idéal ». L'enjeu est donc de déterminer une stratégie de dégradation du signal qui minimise la modification du comportement des observateurs, en tentant de conserver le niveau de visibilité des éléments visuels présentés aux observateurs.

Nous proposons³ une méthodologie d'évaluation spécifique des algorithmes de transposition de luminance qui a un champ d'application plus large, mais qui doit être adaptée en fonction des critères de « qualité » retenus pour toute utilisation particulière des images affichées. Elle est présentée ci-dessous, ainsi que son application dans le cas des scènes routières diurnes.

¹ Cathode Ray Tube (tube cathodique).

² Rapport entre la luminance maximum et la luminance minimum affichable.

³ Travaux réalisés en collaboration avec l'équipe « Vision » du *Muséum National d'Histoire Naturelle*, dans le cadre du projet Predit VOIR.

La performance visuelle est une performance associée à une tâche liée à un stimulus visuel. Il s'agit en général de tâches de détection ou d'identification d'objets dans une scène (marquage, signalisation, véhicule, etc.), pour lesquelles les réponses des observateurs ont une valeur de vérité (vrai ou faux). A partir du taux de bonnes réponses et éventuellement du délai de réponse, il est possible de définir un indicateur objectif de la performance visuelle de l'observateur pour une tâche spécifique (CIE, 2002).

L'élément fondamental de notre approche consiste à utiliser un cadre de référence extérieur au domaine des images de synthèse (Brémond, 2002). Une scène matérielle est construite, dans laquelle des observateurs vont être soumis à une série de tests, permettant d'apprécier leurs performances visuelles pour la tâche considérée. Cette scène de référence est ensuite caractérisée photométriquement et numérisée. L'image numérique qui en résulte, considérée comme une représentation « fidèle » de la scène réelle, est ensuite visualisée sur un moniteur, après avoir été modifiée par un algorithme de transposition de luminances. Les performances visuelles mesurées sur la scène de référence sont alors évaluées sur les images de synthèse visualisées à l'écran. C'est la plus ou moins bonne correspondance entre la valeur de référence de l'indicateur et la valeur obtenue à partir des images de synthèse qui permettra de juger de la qualité de la restitution, *pour l'objectif poursuivi*, de l'ensemble « algorithme de transposition / système de restitution ».

4 Visibilité routière diurne

4.1 Tâche visuelle

Nous nous sommes intéressés avec cette méthodologie au problème de la visibilité routière diurne, et à la pertinence des images de synthèse pour étudier ce problème porteur d'enjeux importants en termes de sécurité routière.

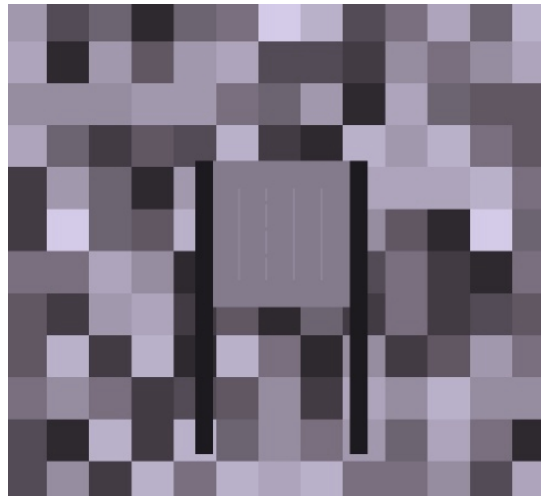


Figure 2 - Modélisation de la scène de référence utilisée pour évaluer les performances visuelles des automobilistes.

Une tâche visuelle a été définie (la discrimination d'une ligne parmi 4 lignes parallèles, Cf. Fig. 2), un moyen de restitution choisi (un moniteur CRT). Une plage de luminance a également été fixée à partir des connaissances sur les niveaux lumineux usuels que rencontre un automobiliste de jour. Il ne s'agissait pas de simuler une situation de conduite, mais de définir une tâche visuelle proche d'une tâche de détection en situation de conduite, du point de vue des paramètres visuels (luminances, contrastes, dispersion, etc.) sans introduire de biais par des « objets » ayant une sémantique, qui ajouterait à la tâche visuelle une composante cognitive parasite.

4.2 Référence

Un système de présentation d'images imprimées a été mis au point pour évaluer les performances visuelles dans ce cadre de référence. Le dispositif est composé d'un mur de polystyrène éclairé par un

« projecteur de transparents », de manière à projeter en périphérie de la zone test un bruit blanc de carrés aléatoirement répartis, et d'un moteur pas-à-pas, piloté par l'observateur, qui permet de présenter des séries de tests. Pour réaliser ces tests, des images correspondant à la tâche visuelle ont été imprimées avec différents niveaux de contraste entre les lignes et le fond, compris entre $-0,022$ et $0,024$. L'environnement lumineux était compris entre 40 et 800 cd/m^2 . Les observateurs, au nombre de 5 , étaient situés à 2 mètres de l'écran.

La tâche visuelle consistait à identifier, parmi 4 lignes, celle qui était différentes des 3 autres (une continue et 3 discontinues, ou le contraire). Le temps de réponse était enregistré, de manière à pouvoir calculer la performance visuelle selon la définition proposée par la CIE, en divisant la proportion de bonnes réponses par le temps de réponse.

4.3 Images de synthèse.

Dans un deuxième temps, cette scène de référence a été caractérisée par des mesures photométriques, d'une part en ce qui concerne l'environnement (qui reste le même d'un test à l'autre), d'autre part en ce qui concerne les tests présentés. A partir de ces mesures de luminance, des images de synthèse en luminance ont été produites. L'écran CRT utilisé pour présenter ces images a été étalonné par des mesures photométriques.

Quatre algorithmes de transposition de luminance ont été sélectionnés dans la littérature (Vénot *et al*, 2002 ; Dumont *et al*, 2003), dont trois sont linéaires. Notre intérêt pour cette dernière catégorie d'algorithmes est lié à leur simplicité, mais surtout au fait qu'ils sont les seuls à introduire un niveau de « distorsion » uniforme du niveau de visibilité (si on fait l'hypothèse que la détection des contrastes est liée au rapport $\Delta L/L$ (LeGrand, 1972)).

L'algorithme 1 (*maximum*) consiste simplement à faire entrer strictement la luminance à afficher dans la dynamique de luminance disponible, ce qui peut s'écrire : $L^d = (L_{\max}^d \times L) / L_{\max}$, si L^d est la luminance affichée, L la luminance à afficher, c'est-à-dire la luminance de la scène de référence au point correspondant, L_{\max}^d la luminance maximum affichable sur le système de visualisation utilisé, et L_{\max} la luminance maximum dans la scène de référence.

L'algorithme 2 (*moyenne*) compense la forte sensibilité du précédent à la valeur d'un seul pixel. Plutôt que de faire correspondre la luminance maximum à afficher avec la luminance maximum affichable, il fait correspondre la luminance moyenne à afficher $\langle L \rangle$ avec la moitié de la luminance maximum affichable : $L^d = (L_{\max}^d \times L) / 2\langle L \rangle$.

L'algorithme 3 (Ward, 1994) introduit la notion de sensibilité de l'œil humain, de manière à essayer de respecter les performances visuelles. La luminance d'adaptation L_a dans la scène de référence est calculée⁴, de même que le seuil de sensibilité correspondant $\Delta L_s(L_a)$. La pente de cette transposition linéaire est calculée de façon à conserver le niveau de visibilité⁵ :

$$L^d = \frac{\Delta L_s(L_{\max}^d / 2)}{\Delta L_s(L_a)} L$$

L'algorithme 4 (Larson *et al*, 1997) n'est pas linéaire. Il conduit qualitativement à de bons résultats dans certaines applications. Il a pour but d'optimiser l'histogramme de luminance, afin d'utiliser au mieux la dynamique de luminances disponibles.

⁴ Dans cet algorithme, la luminance d'adaptation sur écran est supposée égale à la moitié de la luminance maximum affichable.

⁵ ΔL_s est calculé à partir de données expérimentales en sciences de la vision (LeGrand, 1972).

A partir de ces 4 algorithmes, des séries d'images numériques RVB ont été calculées en vue de la visualisation sur l'écran. Nous avons ainsi pu comparer individuellement les performances visuelles des observateurs dans la situation de référence avec leurs performances devant les images résultant de ces 4 algorithmes de *tone mapping*. Cette démarche a ainsi permis d'évaluer la « qualité » relative de ces algorithmes, relativement à une tâche visuelle de discrimination en situation de conduite de jour.

4.4 Résultats.

Dans la situation de référence, nous avons utilisé la notion de « performance visuelle » définie par la CIE. Elle est bonne (proche de 1) pour un contraste élevé, c'est-à-dire que la proportion de bonnes réponses est proche de 100%, et que les réponses sont rapides (de l'ordre de 1 seconde). Elle est proche de 0 pour des contrastes faibles, du fait de la longueur du temps de réponse. La performance visuelle évolue continuellement entre ces extrêmes, ce qui rend cet indicateur particulièrement utile (Fig. 3). La variabilité de cet indicateur d'un observateur à l'autre est une réalité objective, qui doit inciter à éviter les moyennes, et à préférer les comparaisons inter-individuelles.

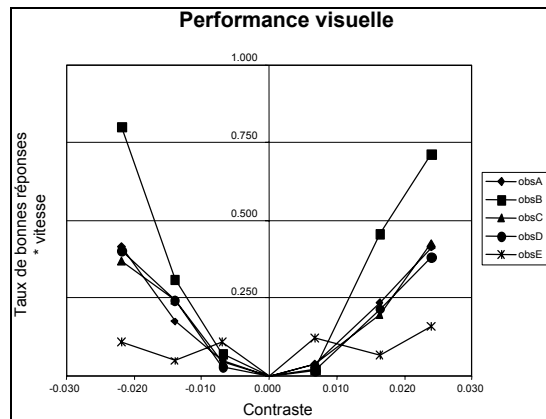


Figure 3 - Performance visuelle en fonction du contraste, pour les 5 observateurs (scène de référence).

Le but de cette étude était de comparer les performances face aux images de synthèse et face à la scène de référence, en partant du principe qu'un algorithme performant devrait conduire à des performances visuelles identiques pour tous les observateurs, et quel que soit le contraste présenté. Malheureusement, aucun des 4 algorithmes testés ne remplit ces conditions (Fig. 4).

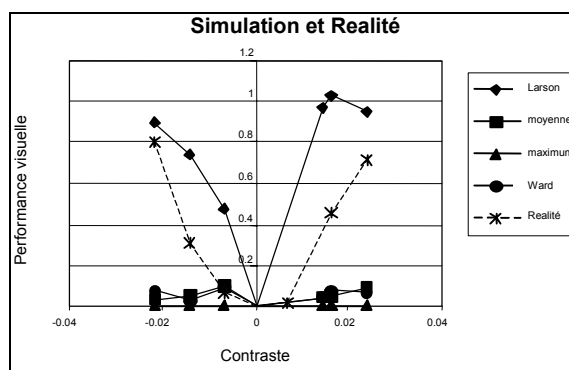


Figure 4 - Comparaison des performances visuelles entre simulations et référence, pour un observateur.

Les algorithmes linéaires retenus conduisent à des contrastes mesurés à l'écran qui varie entre -0,016 et +0,016, alors que l'algorithme de *Larson* conduit à des contrastes qui varient entre -0,06 et 0,11. En conséquence, les 3 algorithmes linéaires conduisent les observateurs à une performance significativement plus faible que la référence (Fig. 4), alors que l'algorithme de *Larson* conduit à des performances significativement plus élevées, sauf pour le contraste le plus faible dans la référence, qui était visualisé à l'écran avec un contraste nul en raison de la quantification.

Finalement, aucun des algorithmes testés ne permet de reproduire les performances visuelles des observateurs de manière satisfaisante, pas même l'algorithme proposé par Ward, pourtant basé sur des données psycho-physiques. Mais cette expérimentation permet de penser que de nouveaux algorithmes, intermédiaires entre les deux classes de performances étudiées ici, pourraient donner de meilleurs résultats.

5 Conclusion

La méthodologie présentée ci-dessus a un domaine d'application relativement large dans le domaine de la synthèse d'images, pour tout ce qui touche au « réalisme ». Elle permet en particulier de lever l'ambiguïté sur ce mot, en quantifiant la notion de réalisme par rapport à un objectif bien défini en termes de perception visuelle (Ferwerda, 2003).

A partir de la méthodologie développée et des résultats obtenus, notre objectif est d'une part de rechercher un algorithme conduisant à des performances intermédiaires entre les deux familles d'algorithmes mises en évidence, et d'autre part d'élargir nos recherches à d'autres situations de conduite, pour lesquelles les enjeux de visibilité routière sont importants : conduite de nuit, conduite en tunnel. Enfin, la méthodologie d'évaluation devra être appliquée pour les différentes solutions technologiques d'affichages utilisées (CRT, LCD⁶, DLP⁷, etc.), puisque l'algorithme le plus performant dépend des performances spécifiques du système de visualisation.

6 Références

Brémond, R. (2002) Respect des niveaux de visibilité dans la restitution d'images de synthèse en unités physiques. *Actes AFIG '02 (Lyon, décembre 2002)*.

CIE Publication 122 (1999) *The relationship between digital and colorimetric data for controlled CRT displays*.

CIE Publication 145 (2002) *The correlation of models for vision and visual performance*.

Dumont, E. (2002) *Caractérisation, modélisation et simulation des effets visuels du brouillard pour l'utilisateur de la route*. Thèse de Doctorat de l'Université Paris V, novembre 2002.

Dumont, E., Brémond, R., Boust, C., da Costa, E., Viénot, F. (2003) Assessment of the visual quality of images for visibility experiments : psychometric evaluation of tone mapping algorithms. *Proceedings 25th Session of the CIE (San Diego, June 2003)*, CIE Publication 152-2003, Vol.2, D4.82-85.

Ferwerda, J.A. (2003) Three varieties of realism in computer graphics. *Proceedings SPIE Human Vision and Electronic Imaging '03*, in press.

LeGrand, Y. (1972) *Optique physiologique – Tome 2 : Lumière et couleurs*. Masson et Cie, Paris, France (2ème édition).

Larson, G.W., Rushmeier, H., Piatko, C. (1997) A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes. *IEEE Transactions on Visualization and Computer Graphics*, 3(4), 291–306.

Tumblin, J., Rushmeier, H. (1993) Tone reproduction for Realistic Images. *IEEE Computer Graphics & Applications*, 13(6), 42–48.

Viénot, F., Boust, C., DaCosta, E., Brémond, R., Dumont, E. (2002) Psychometric assessment of the look and feel of digital images. *Proceedings 7th Driving Simulation Conference 2002 (Paris, September 2002)*, 301-311.

Ward, G. (1994) A Contrast-based Scale Factor for Luminance Display. *Graphics Gems IV*, ed. P. S. Heckbert, 391–397.

⁶ Liquid Crystal Display : technologie d'affichage à cristaux liquides.

⁷ Digital Light Processor : technologie d'affichage à micro-miroirs.