

# Conception et évaluation d'un opérateur de reproduction de tons pour des études de visibilité routière

J. Grave

Laboratoires LCPC / LIRIS

R. Brémond

Laboratoires LCPC

justine.grave, roland.bremond@lcpc.fr

**Résumé :** *Aujourd'hui les dispositifs d'affichage ne nous permettent d'afficher qu'un domaine restreint de luminances. Dans ce papier, nous présentons un opérateur de reproduction de tons (TM), qui a pour fonction de convertir les images à forte dynamique de luminances (HDR) en images pouvant être affichées. Cet algorithme a été conçu pour des études de visibilité routière avec pour objectif de préserver la perception visuelle des observateurs pour une tâche visuelle de détection. Nous reprenons l'idée que le système visuel humain fonctionne comme un ensemble de filtres passe-bande et que la fonction de sensibilité au contraste serait l'enveloppe de leur réponse. L'image HDR est d'abord décomposée en une pyramide Laplacienne qui fournit une pyramide d'images de contrastes. Chaque niveau est ensuite traité séparément de manière à préserver la perception des contrastes. En visibilité routière, on utilise le niveau de visibilité pour quantifier cette perception. Nous proposons ensuite une expérience psychophysique pour évaluer notre algorithme. Nous comparons les performances visuelles d'un observateur face à une scène de référence et face à une image de cette scène de référence traitée par un opérateur TM et affichée sur un écran LCD calibré. La scène de référence est une image HDR projetée par un projecteur DLP calibré. Le DLP a une luminance maximale affichable de  $500 \text{ cd.m}^{-2}$ , avec une dynamique de  $1000 : 1$ , alors que la luminance maximale affichable par le LCD est de  $167 \text{ cd.m}^{-2}$ , avec une dynamique de  $400 : 1$ . Nous mesurons les performances visuelles des observateurs avec un anneau de Landolt. Les mesures sont effectuées pour 8 valeurs de contraste entre l'anneau et le fond. Une brisure apparaît pendant 100 ms et nous demandons aux sujets d'indiquer sa position. Ces tests ont permis de mettre en évidence que notre algorithme offre un meilleur rendu des contrastes autour du seuil de perception, pour une scène de conduite diurne, que 3 autres opérateurs utilisés pour la comparaison.*

**Mots-clés :** Tone Mapping, Visibilité routière, Performances visuelles, Luminance.

## 1 Introduction

Les études de visibilité routière en situation de conduite utilisant des images de synthèse offrent de nombreux avantages par rapport aux essais sur piste ou sur route (reproductibilité des essais, absence de risque). Mais la richesse de l'environnement visuel de l'utilisateur de la route dépasse largement les capacités matérielles des dispositifs de visualisation (dynamique de luminance, domaine de couleur, quantification, résolution). Pour contourner ces contraintes, on propose un algorithme de reproduction de tons (Tone Mapping) qui transpose, en utilisant les caractéristiques de la perception visuelle humaine, une image à forte dynamique de luminances (HDR), calculées et/ou mesurées, en une image à luminances affichables. Après le choix d'une tâche visuelle pertinente par rapport à la tâche de conduite, des tests psychophysiques sont élaborés dans le but d'évaluer cet algorithme. Ces tests visent à comparer les performances visuelles basées sur l'acuité visuelle et la sensibilité au contraste dans une scène de référence (une image HDR) et dans cette même scène dont la dynamique des luminances a été compressée grâce à l'opérateur TM.

## 2 Etat de l'art

Le problème de la reproduction des luminances sur un dispositif d'affichage limité en niveaux de luminance ou en dynamique n'est pas récent. Les premiers à s'être intéressés à cette question sont les photographes qui traitaient les photos de manière à ce qu'elles correspondent plus à l'apparence visuelle des scènes photographiées.

Pour les images de synthèses de nombreux opérateurs TM ont été proposés [DCWP02]. Il existe deux catégories d'opérateurs. La première regroupe les opérateurs globaux ou spatialement uniforme. L'image est considérée dans son ensemble et tous les pixels subissent la même transformation. La seconde regroupe les opérateurs locaux qui appliquent des transformations différentes aux différentes parties de l'image.

Parmi les opérateurs globaux, on peut mentionner celui de Ward [War94] qui a une base théorique psycho-visuelle. C'est un opérateur linéaire, très simple, qui calcule un facteur d'échelle grâce au modèle de performance visuelle de Blackwell [CIE81]. Il a pour fonction de convertir de la même manière (par le même facteur), de la scène réelle à l'image affichée, les luminances et le seuil de différence discernable. Cette approche est utile pour visualiser des scènes où l'analyse de la visibilité est primordiale. On peut également citer l'algorithme de Larson et al. [LRP97]. Il effectue une compression des intervalles de luminances sous-représentées sans dilatation des intervalles de luminances sur-représentées, par ajustement d'histogramme. L'opérateur de Pattanaik et al. [PFFG98] fait parti de la seconde catégorie. Il utilise un modèle de vision rigoureux pour la sensibilité au contraste et traite l'image en couleur, en fonction de ses fréquences spatiales. Les modèles de vision qu'il utilise sont à la fois des modèles de performance et d'apparence visuelle. Celui de Reinhard et al. [RSSF98] s'inspire des techniques utilisées par la photographie traditionnelle en jouant sur les temps d'exposition pour différentes zones de l'image. Des opérateurs proposent également une solution dynamique [PTYG98].

Des études ont également déjà été proposées pour évaluer des opérateurs de reproduction de tons. Drago et al. [DMMS03] demandaient aux sujets de juger des images traitées pour un opérateur TM. Ils ont déterminé un stimulus de référence qu'ils ont utilisé pour déterminer l'opérateur qui fournissait la réponse la plus proche de cette référence, d'un point de vue perceptuel. D'autres études ont été effectuées, où les observateurs indiquaient leur préférences parmi les images traitées par rapport à une scène réelle. Ledda et al. [LCTS05] ont proposé des expériences utilisant une image HDR comme scène de référence et ont demandé aux sujets de comparer les opérateurs deux à deux. Enfin, Viénot et al. [VBC<sup>+</sup>02] présentent une méthode dévaluation basée sur l'apparence visuelle.

Dans cet article, nous cherchons d'une part un opérateur de reproduction de tons local, dépendant des fréquences spatiales de l'image, traitant l'image en luminance et se focalisant sur la préservation des performances visuelles de l'observateur. Ce qui, pour une première approche, est primordial pour des études de visibilité routière. D'autre part nous souhaitons évaluer cet opérateur à l'aide d'expériences psychophysiques qui utilisent un stimulus qui corresponde à une tâche visuelle de détection.

### 3 L'opérateur de restitution d'images

#### 3.1 Modèle de vision

L'idée de départ (suggérée par Campbell et Robson en 1968 [CR68]), est que le système visuel se compose de groupes de filtres passe-bande indépendants et quasi-linéaires, de bandes de fréquences plus étroites que celle de la fonction de sensibilité au contraste. Celle-ci serait donc l'enveloppe de la sensibilité de multiples filtres, ou canaux. Le principe est donc de décomposer l'image en une pyramide d'images, chaque image représentant des bandes de fréquences différentes de l'image de départ. Pour cette décomposition nous utilisons, comme Pattanaik et al. [PFFG98], la méthode de la pyramide laplacienne de Burt&Adelson [BA83]. Un filtre gaussien est appliqué à chaque niveau de la pyramide (l'image filtrée pour le niveau  $l$  est notée  $L_w^l$ ) pour obtenir l'image du niveau suivant. La différence des images de deux niveaux consécutifs fournit une image "contraste". Suit une étape de traitement d'image qui consiste à appliquer des modèles de vision à ces images "contraste" (cf. figure 1).

L'image du niveau  $l - 1$  est d'abord filtrée, par un filtre gaussien  $5*5$  (cf tableau 1), qui réduit la taille de l'image par 4 :

$$L_w^l(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) L_w^{l-1}(2i + m, 2j + n) \quad (3.1)$$

L'image du niveau  $l$ ,  $L_w^l(i, j)$ , avec  $w$  pour "world", est ensuite agrandie, par interpolation bilinéaire, pour revenir à la taille du niveau  $l - 1$ , elle est alors notée  $L_{wExpand}^l(i, j)$ , puis l'image contraste  $C_w^l(i, j)$  est obtenue par la différence :

$$C_w^l(i, j) = L_w^{l-1}(i, j) - L_{wExpand}^l(i, j) \quad (3.2)$$

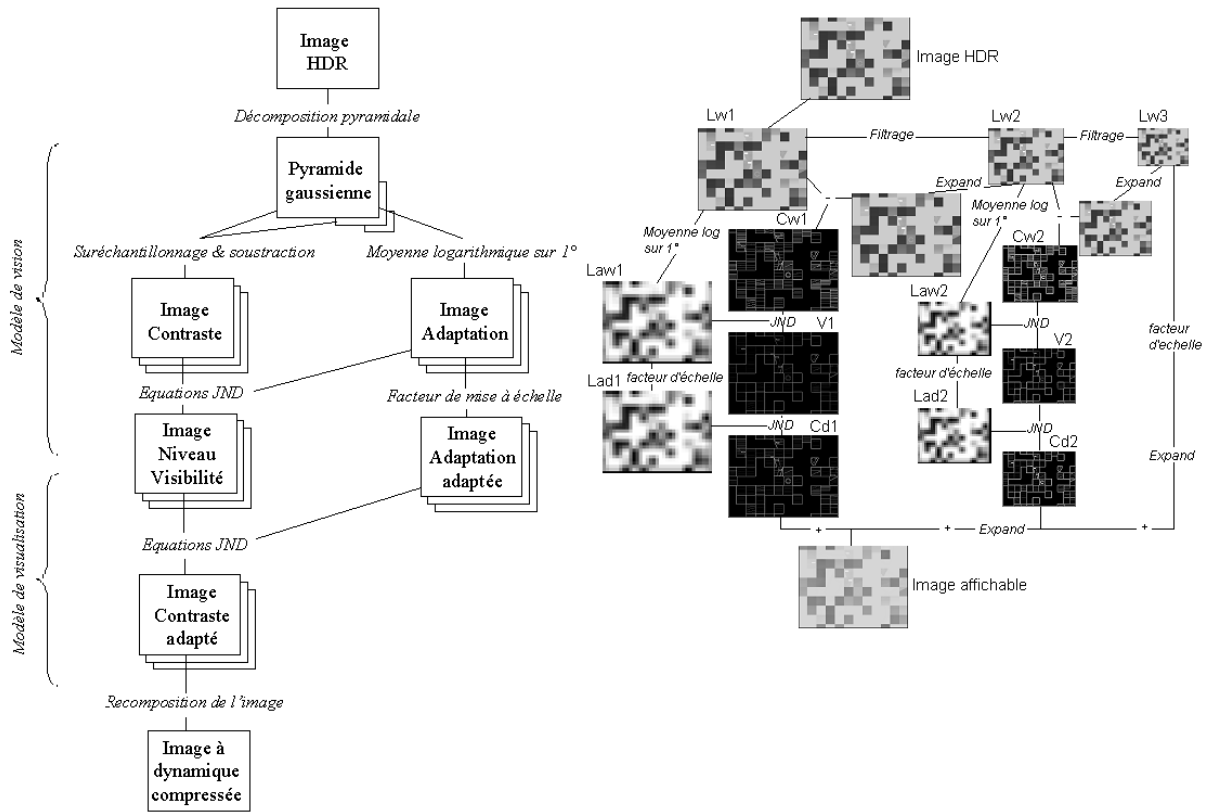


FIG. 1 – Algorithme de l’opérateur de restitution d’image. A droite, 3 niveaux de la pyramide sont illustrés (Il faut une pyramide à 7 niveaux pour couvrir tout le domaine de sensibilité du système visuel)

0.0025	0.0125	0.02	0.0125	0.0025
0.0125	0.0625	0.1	0.0625	0.0125
0.02	0.1	0.16	0.1	0.02
0.0125	0.0625	0.1	0.0625	0.0125
0.0025	0.0125	0.02	0.0125	0.0025

TAB. 1 – Filtre gaussien  $5 * 5$ ,  $w$ , utilisé pour la décomposition pyramidale [BA83]

Notre objectif est de conserver la perception des contrastes. C’est à dire que l’observateur perçoit de la même manière les contrastes de l’image HDR et ceux de l’image affichée. Pour mesurer la perception des contrastes du point de vue des performances visuelles, on utilise le niveau de visibilité (VL) décrit, pour la visibilité routière [Adr91], par :

$$V^l(i, j) = \frac{C_w^l(i, j)}{\Delta L_s(L_{aw}^l(i, j))} \quad (3.3)$$

où  $V^l$  est l’image des VL pour le niveau  $l$  de la pyramide,  $\Delta L_s$  la différence de luminances au seuil de la perception et qui dépend de la luminance d’adaptation  $L_{aw}^l$ .

On cherche donc à ce que le VL à chaque pixel de l’image soit conservé par l’opérateur. On utilise le système d’équations *Just Noticeable Difference* (JND) décrit par Larson et al. [LRP97] (dont les données proviennent des études de Ferwerda et al. [FPSG96]) pour calculer les valeurs de contraste seuil  $\Delta L_s$  en fonction de la luminance d’adaptation (LA), notée  $L_{aw}^l(i, j)$ . Dans notre cas, les valeurs de LA sont obtenues à partir de l’image  $L_{wExpand}^{l-1}(i, j)$  en calculant une moyenne logarithmique des luminances sur des blocs de 1 degré d’envergure. On considère en effet que l’adaptation visuelle se fait principalement en vision fovéale, qui a un champ de l’ordre de 1 degré [Gra]. Pour accélérer les calculs, on sous-échantillonne l’image de LA en calculant une valeur tous les

1/2 degré. La valeur de LA pour le pixel  $(i, j)$ , notée  $L_{aw}^l(i, j)$ , est calculée à partir des 4 valeurs de LA les plus proches par interpolation bilinéaire.

### 3.2 Modèle de visualisation

Les images en VL,  $V^l(i, j)$ , sont utilisées pour recalculer de nouvelles images contrastées  $C_d^l(i, j)$  adaptés à la dynamique de l'écran, sur lequel sera affichée l'image reconstruite, avec de nouvelles valeurs de luminances d'adaptation  $L_{ad}^l(i, j)$ , avec  $d$  pour "display" :

$$C_d^l(i, j) = V^l(i, j) \cdot \Delta L_s(L_{ad}^l(i, j)) \quad (3.4)$$

Les nouvelles luminances d'adaptation  $L_{ad}^l(i, j)$  sont calculées en multipliant les anciennes valeurs  $L_{aw}^l(i, j)$  par le facteur  $k$  [War94], dont on impose une valeur inférieure ou égale à 1 :

$$L_{ad}^l(i, j) = k \cdot L_{aw}^l(i, j) \quad (3.5)$$

$$k = \left[ \frac{1,219 + (L_{dmax}/2)^{0,4}}{1,219 + (L_{wa}^l)^{0,4}} \right]^{2,5} \quad (3.6)$$

où  $L_{dmax}$  est la luminance maximale affichable par l'écran et  $L_{wa}^l$ , la moyenne logarithmique des luminances de l'image  $L_{wa}^l(i, j)$ .

Enfin l'image est reconstruite, suivant le principe inverse de la décomposition pyramidale. Le dernier niveau de la pyramide subit un traitement différent puisqu'il ne s'agit pas d'une image "contraste" mais d'une image des basses fréquences. On lui applique d'abord le facteur  $k$ . Elle est ensuite agrandie par interpolation bilinéaire et ajoutée à l'image "contraste" du niveau inférieur. L'image obtenue est à son tour agrandie par interpolation bilinéaire et ajoutée à l'image du niveau inférieur et ainsi de suite avec tous les niveaux de la pyramide. L'image finale est une image de luminances affichables (cf. le schéma de l'algorithme figure ??). La figure 2 présente l'effet de différents opérateurs de TM sur l'histogramme d'une image utilisée pour l'évaluation. La figure 3 illustre les effets des opérateurs de Ward [War94], Larson et al. [LRP97], Pattanaik et al. [PFFG98] et celui proposé ici. Il n'est pas possible d'évaluer ces opérateurs à partir des images imprimées ici, qui ont été calculées pour être affichées sur un écran LCD spécifique.

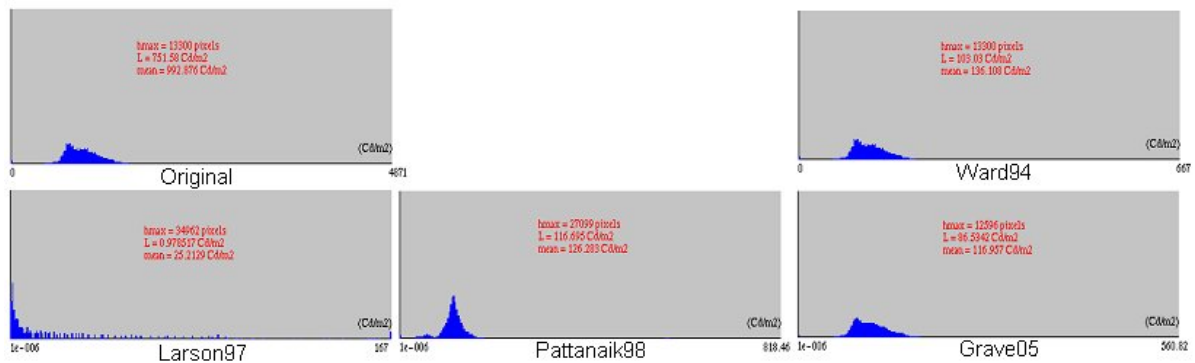


FIG. 2 – Histogrammes de l'image HDR originale et des images traitées par différents TM

A ce stade des recherches, il est important d'évaluer cet algorithme, associé à un dispositif d'affichage, pour pouvoir, le cas échéant, le valider, le modifier et/ou le compléter.

## 4 L'évaluation

Le principe général des expérimentations qui permettront l'évaluation du dispositif de visualisation [GB05] est de comparer les performances visuelles de l'observateur lorsqu'il accomplit la tâche visuelle à partir d'une scène

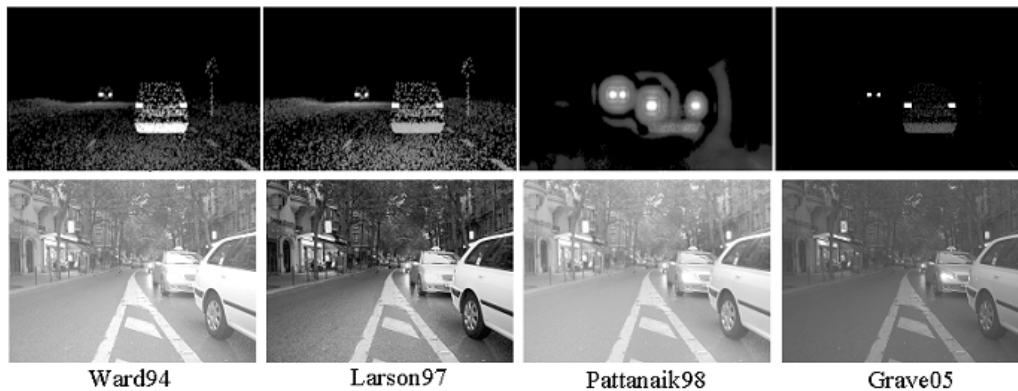


FIG. 3 – La première série est une image de synthèse d’une scène routière nocturne, traitée par 4 opérateurs de TM,  $L_{max} = 16180.9 \text{ Cd.m}^{-2}$  et  $L_{min} = 0 \text{ Cd.m}^{-2}$ . La seconde série est une image mesurée d’une scène urbaine en soirée, traitée par 4 opérateurs de TM avec  $L_{max} = 478 \text{ Cd.m}^{-2}$  et  $L_{min} = 0 \text{ Cd.m}^{-2}$ .

de référence (ici une image HDR) et lorsqu’on lui présente la scène à laquelle on applique un opérateur TM (cf. le protocole de Viénot et al. [VBC<sup>+</sup>02]). Ces différences de performances entre la scène de référence et la scène traitée sont mesurées puis comparées pour différents opérateurs de restitution, celui proposé et ceux de [War94], de [LRP97] et de [PFFG98] (cf. figure 4). Les tests ont lieu au LCPC<sup>1</sup> dans une salle fermée, sans fenêtre et les murs peints en noir. Cela permet le contrôle de la photométrie et la répétabilité des conditions expérimentales.

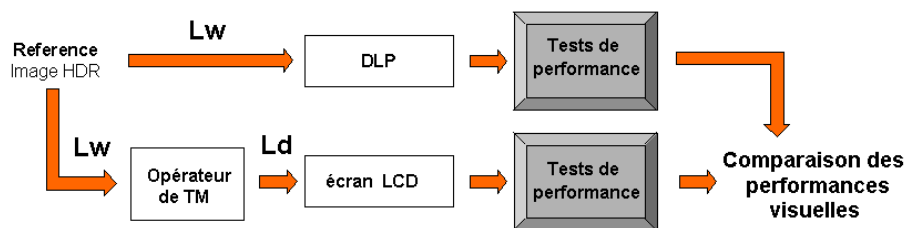


FIG. 4 – Principe général de l’évaluation de l’opérateur de restitution

## 4.1 Test de performance

On a retenu la détection d’un danger comme tâche visuelle pertinente pour la conduite. Cette tâche est associée à des indices de performances. On a choisi de tester les performances visuelles de l’observateur à l’aide d’un anneau de Landolt. Cela constitue une simplification forte de la tâche visuelle qui nous intéresse mais qui donne accès à un aspect fondamental pour cette tâche visuelle et qui est très souvent utilisé dans les études de visibilité routière [Adr03].

### 4.1.1 L’anneau de Landolt

L’anneau brisé de Landolt est le plus simple des optotypes<sup>2</sup>. Cette forme est déjà complexe et la détection de la brisure dépend de plusieurs paramètres (cf. figure 5) :

- la taille de la brisure  $e$  (en minutes d’angle). L’usage est de caractériser la brisure non plus par  $e$  mais par l’acuité visuelle définie par  $V = \frac{1}{e}$
- la position de la brisure (qui peut être présentée dans 4 ou 8 positions différentes),
- la luminance du fond  $L_b$ ,

<sup>1</sup>salle CLOVIS : Clovis est un LOcal de Visualisation d’Images de Synthèse

<sup>2</sup>Un optotype est un test visuel qui comporte, jusqu’à un certain point, la reconnaissance d’une forme.

- le contraste (défini comme la fraction de Weber) entre la luminance du fond  $L_b$  et la luminance de l'anneau  $L_t$ , défini par  $C = \frac{L_b - L_t}{L_b} = \frac{\Delta L}{L_b}$
- le temps de présentation de la brisure  $t$ .

L'anneau est présenté sur un carré de luminance uniforme (l'ensemble constitue ce qu'on appellera le test). On souhaite ajouter un bruit visuel (cf. paragraphe 4.1.2). Pour intégrer correctement le test dans l'image de bruit, on choisit la luminance du carré égale à la luminance d'adaptation de l'image de bruit.



FIG. 5 – Description de l'anneau de Landolt pour le test de performance

Dans le choix des tests, on fixe les paramètres suivant :

- le temps de présentation de la brisure est fixé à  $100ms$ ,
- la luminance du carré  $L_b$  est égale à la luminance d'adaptation calculée selon [MS45] avec la position du test dans l'image comme centre du regard,
- l'acuité visuelle est fixée à  $V = 0.13$  (calculée suivant la géométrie de la scène et la position de l'observateur) qui correspond à  $e = 8$  minutes d'angle environ.

Il reste donc deux paramètres qui varient :

- le contraste. On choisit plusieurs valeurs de contraste allant d'une valeur proche du seuil de détection à une valeur détectable sans difficulté (8 valeurs de contraste),
- la position de la brisure, qui peut avoir 4 positions différentes (cf. figure 5).

#### 4.1.2 Bruit visuel

Le test de référence est un anneau de Landolt présenté sur un carré de luminance uniforme. L'image dans laquelle est incrusté le test est choisie selon plusieurs critères :

1. un histogramme des luminances très proche de celui d'une scène de conduite dans des conditions de visibilité déterminées (brouillard de jour ou de nuit, conduite de nuit, etc.). Dans l'étude présentée ici, dont on montre les résultats en section 5, il s'agit d'une scène de brouillard de jour ([Dum04], cf figure 6),
2. une absence de sémantique pour limiter l'aspect cognitif de l'environnement visuel et ne pas biaiser les performances par des facteurs sémantiques spécifiques à une scène particulière.

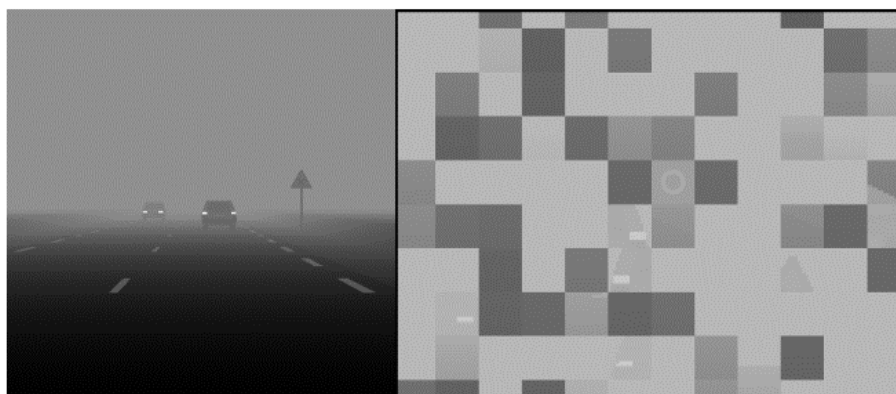


FIG. 6 – Image utilisée pour fabriquer l'image de bruit visuel dans laquelle est incrustée le test (à gauche) et l'image de test avec l'anneau au centre (à droite)

Pour fabriquer cette image de bruit, une image de scène routière en luminances simulées ([Dum04]) et à forte dynamique (incluant des sources lumineuses) est seuillée à  $500Cd.m^{-2}$  (puisque c'est la luminance maximale

affichable par le vidéoprojecteur qui affichera l'image HDR). Elle est ensuite découpée en blocs de la même taille que le test (pour supprimer la sémantique). Ces blocs sont mélangés aléatoirement pour former une nouvelle image ayant le même histogramme que l'image de départ, mais sans la sémantique.

## 4.2 Protocole expérimental

On a choisi de comparer les performances de l'observateur face à une scène de référence, qui offre une dynamique de luminance importante (1000/1 avec une luminance affichable maximale de  $500cd.m^{-2}$ ), et face à la même scène traitée par un opérateur de TM et affichée sur un écran à dynamique restreinte (400/1 avec une luminance affichable maximale de  $167cd.m^{-2}$ ).

### 4.2.1 La scène de référence

La scène de référence nécessite deux dispositifs de projection d'image, pour nous permettre d'afficher des luminances relativement élevées. Une première image est projetée par un vidéoprojecteur DLP (Digital Light Processing)<sup>3</sup>, préalablement calibré selon la méthode du rapport 122 de la CIE [CIE96], au centre de l'écran. Les avantages de ce vidéoprojecteur sont :

1. l'affichage de luminances peut atteindre  $500cd.m^{-2}$  dans les conditions géométriques retenues (image sur l'écran d'environ  $55cm \times 70cm$ ),
2. le pilotage de l'affichage des images (scène de référence numérique) et l'enregistrement des réponses de l'observateur sont facilités par un logiciel d'expérimentation psychophysique<sup>4</sup> installé sur un ordinateur. Le protocole est proche de celui de scène de Viénot et al. [VBC<sup>+</sup>02] mais la référence qu'ils utilisent est physique et donc nécessite un pilotage plus complexe sans pour autant permettre d'avoir des luminances plus fortes.
3. la possibilité de tester ultérieurement les performances d'un observateur dans la détection de mouvement, par l'affichage de courtes séquences vidéo.

Une seconde image est projetée par un rétroprojecteur<sup>5</sup> pour la vision périphérique (d'environ  $1,5m$  par  $1,5m$  pour la scène de référence et 4 fois plus petite pour la configuration avec l'image affichée par le LCD) (cf. figure 7).

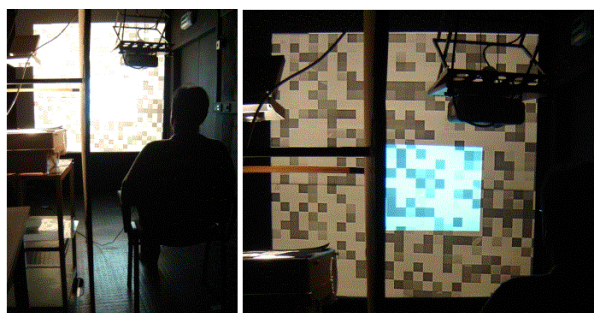


FIG. 7 – Configuration pour la scène de référence dans la salle CLOVIS

### 4.2.2 La scène de comparaison

La configuration que l'on choisit pour la scène de comparaison est très proche de celle de la scène de référence. L'image projetée par le DLP est traitée en luminances par un opérateur de TM que l'on souhaite évaluer (celui de Ward, celui de Larson ou al., celui de Pattanaik et al. et celui qui est proposé ici). L'image est ainsi affichable

<sup>3</sup>Vidéo projecteur DLP NEC LT 1065, 2100 ANSI lumen

<sup>4</sup>logiciel Présentation

<sup>5</sup>Rétroprojecteur PROLITE 250/400, puissance max 300W.

sur un écran LCD <sup>6</sup>, qui a été préalablement calibré avec le même mode opératoire que pour le DLP. L'image du rétroprojecteur est atténuée par un filtre neutre et projetée sur un écran de polystyrène dans lequel s'incruste l'écran LCD. Le filtrage permet d'assurer une diminution des luminances projetées pour la vision périphérique, comparable à celle subie par les luminances des images affichées sur le LCD.

### 4.2.3 Le protocole

Pour chaque test, une image, avec un anneau plein et l'un des 8 contrastes, est affichée. Au bout de 1000 ms une brisure apparaît dans l'anneau durant 100 ms, dans l'une des quatre positions possibles. L'observateur indique à l'aide des touches d'un gamepad la position à laquelle il a vu la brisure. On exige une réponse pour chaque test, même si l'observateur pense n'avoir rien vu (il répond alors au hasard). Après sa réponse, le contraste suivant est affiché pour un nouveau test.

### 4.2.4 L'observateur

On a effectué l'évaluation avec 9 observateurs différents, 3 femmes et 6 hommes. 5 sont âgés entre 20 et 35 ans et 4 entre 35 et 50 ans. Chaque observateur est placé à 3 m de la scène de référence et à 1.5 m de l'écran LCD (pour que l'image soit vue avec la même taille angulaire que l'image du DLP). On lui présente 136 tests pour la scène de référence et pour chacun des opérateurs TM (17 tests par contraste). Les tests durent de 15 à 20 minutes dans chaque configuration. Cette expérimentation se déroule en deux temps. On procède un jour au test avec la scène de référence, et un autre jour avec les scènes traitées par un opérateur TM.

## 5 Résultats

Les données obtenues avec cette expérience ne permettent pas une analyse significative pour chaque observateur. On considère un observateur moyen dont les performances visuelles sont montrées figure 8. On peut comparer les performances visuelles obtenues avec la scène de référence (image HDR) et celles obtenues avec l'image traitée par notre opérateur. Pour permettre une meilleure comparaison, on calcule une estimation de la marge d'erreur autour des valeurs de la référence, pour chaque contraste. On considère les différents tests, pour un contraste  $c$  donné, comme des réalisations d'une variable aléatoire dont l'espérance est la performance de l'observateur moyen  $p_c$ . Soit la variable aléatoire  $X_c$ , qui suit la loi de Bernouilli. Elle ne peut prendre que deux valeurs : 1 (*succès*), avec la probabilité  $p_c$ , ou 0 (*échec*), avec la probabilité  $q_c = 1 - p_c$ . Soit  $Y_c$ , la variable aléatoire correspondant au nombre de bonnes réponses parmi les 153 (17 tests pour chacun des 9 observateurs) pour le contraste  $c$ .  $Y_c$  suit une loi binomiale de paramètres  $n = 153$  et  $p_c$ . L'inégalité de Bienaymé-Tchebychev nous donne :

$$P(|E(Y_c) - Y_c| \geq \epsilon_c) \leq \frac{V(Y_c)}{\epsilon_c^2}$$

Si on note  $\alpha = \frac{V(Y_c)}{\epsilon_c^2}$ , alors  $\epsilon_c = \sqrt{\frac{V(Y_c)}{\alpha}}$ . On note  $F_c^n = E(Y_c)/n$ , la fréquence de bonnes réponses. On sait que  $E(Y_c) = n.p_c$  et  $V(Y_c) = n.p_c.(1 - p_c)$  (propriétés de la loi binomiale). Donc  $E(F_c^n) = p_c$  et  $V(F_c^n) = p_c.(1 - p_c)/n$ . On cherche alors  $\epsilon_c = \sqrt{\frac{p_c.(1-p_c)}{n.\alpha}}$  tel que :

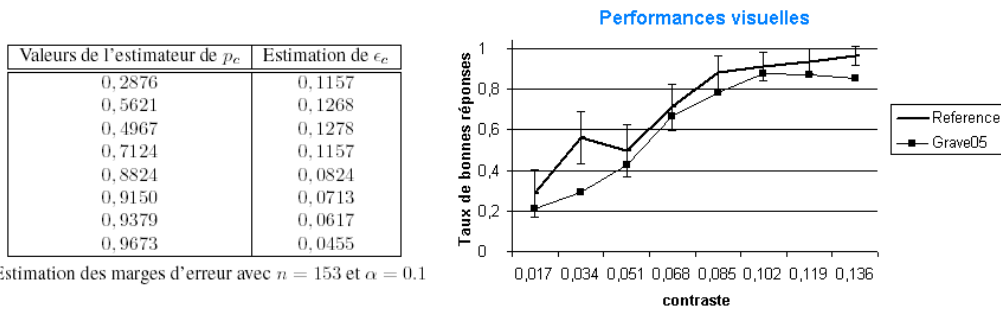
$$\begin{aligned} P(|F_c^n - p_c| \geq \epsilon_c) &\leq \alpha \\ \Leftrightarrow P((F_c^n - \epsilon_c) \leq p_c \leq (F_c^n + \epsilon_c)) &\geq (1 - \alpha) \end{aligned} \quad (5.1)$$

Si on choisit  $\alpha = 0,1$ , cela revient à chercher la marge d'erreur (+ ou -  $\epsilon_c$  autour de la valeur observée  $F_c^n$ ) qui contient 90 pourcents des valeurs pouvant être prise par la variable  $F_c^n$ . On estime ainsi les valeurs de  $\epsilon_c$  par  $\sqrt{\frac{p_c.(1-p_c)}{n.0,1}}$  (cf figure 8).

La figure 8 illustre les performances visuelles mesurées avec la référence et notre opérateur et les marges d'erreur estimées pour la référence. Les valeurs obtenues avec notre opérateur sont pour la plupart dans la marge d'erreur, à l'exception du contraste 0.034 où c'est la valeur de la référence qui semble aberrante.

<sup>6</sup>écran LCD 1701 NEC, 17 pouces





Estimation des marges d'erreur avec  $n = 153$  et  $\alpha = 0.1$

FIG. 8 – Performances visuelles (taux de bonnes réponses) pour l'observateur moyen

## 6 Discussion

Les résultats des tests, dans une situation représentant, du point de vue de la distribution des luminances, une scène routière de brouillard de jour, ont montré que l'algorithme de TM que nous proposons permet de meilleurs résultats que ceux de Ward, de Larson et al. et de Pattanaik et al., du point de vue des performances visuelles, pour des valeurs du contraste autour du seuil de détection (cf. figure 9).

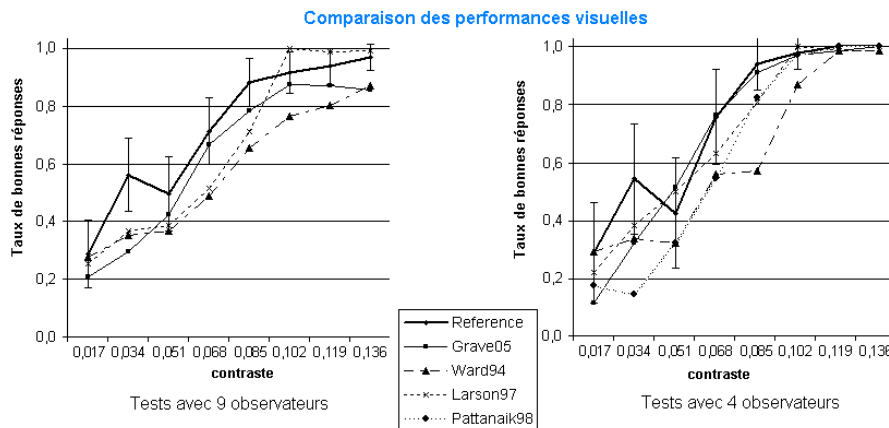


FIG. 9 – Comparaison des performances visuelles obtenues pour 9 opérateurs dans le cas des courbes de gauche et 4 observateurs dans le cas des courbes de droite, pour 4 opérateurs TM

## 7 Conclusion et perspectives

Nous avons mis en place un nouvel algorithme de reproduction de tons, adapté à une tâche visuelle de détection, dans le but de réaliser des études de visibilité routière. Notre objectif principal était de préserver les performances visuelles de l'observateur malgré la compression de la dynamique de luminances des images. Pour évaluer notre algorithme et vérifier qu'il remplissait au mieux cette condition, nous avons mis au point une expérience psychophysique qui permettait de comparer les performances visuelles des observateurs lorsqu'ils accomplissaient une tâche de détection avec une image HDR et avec la même image, traitée par un opérateur TM et affichée sur un écran à dynamique de luminances restreinte. Les résultats des expériences menées montrent que l'opérateur que nous proposons permet de mieux préserver les performances visuelles, pour une scène de conduite diurne, autour du seuil de perception, que les autres opérateurs utilisés pour l'évaluation.

Nous envisageons élargir ces expériences à des images représentant d'autres conditions de conduite (urbaine, nocturne, en soirée) et ajouter dans notre évaluation l'opérateur proposé par Reinhard et al. [RSSF98] qui est apparu dans l'évaluation de Ledda et al. [LCTS05] comme un opérateur performant malgré son absence de modèle

de vision. Nous prévoyons également effectuer des tests avec un dispositif d'affichage de simulateur de conduite <sup>7</sup>, beaucoup plus contraignant, du point de vue de la dynamique des luminances, qu'un écran LCD.

## Références

- [Adr91] W. Adrian. Visibilité de cibles, modèle de calcul. *Lux*, 161, 1991.
- [Adr03] W. Adrian. The effect of observation time and contrast on visual acuity. *Clinical and experimental optometry*, (86) :179–182, 2003.
- [BA83] P. J. Burt and E. H. Adelson. The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4) :532–540, April 1983.
- [CIE81] CIE. An analytic model for describing the influence of lighting parameters upon visual performance. 1(3), 1981.
- [CIE96] CIE. The relationship between digital and colorimetric data for computer-controlled crt displays. 1996.
- [CR68] F.W. Campbell and J.G. Robson. Application of fourier analysis to the visibility of gratings. *Journal of Physiology*, (197) :551–566, 1968.
- [DCWP02] K. Devlin, A. Chalmers, A. Wilkie, and W. Purgathofer. Tone reproduction and physically based spectral rendering. *State of the Art Reports, Eurographics 2002*, pages 101–123, 2002.
- [DMMS03] F. Drago, W. L. Martens, K. Myszkowski, and H.-P. Seidel. Perceptual evaluation of tone mapping operators. *ACM SIGGRAPH Conference Abstracts and Applications*, 2003.
- [Dum04] E. Dumont. Simulation photométrique des effets du brouillard sur la vision. *RFIA 2004*, pages 53–60, 2004.
- [FPSG96] J.A. Ferwerda, S. Pattanaik, P.S. Shirley, and D.P. Greenberg. A model of visual adaptation for realistic image synthesis. *Proc. SIGGRAPH '96*, pages 249–258, 1996.
- [GB05] J. Grave and R. Brémond. Designing a tone mapping algorithm for road visibility experiments. *Poster abstract in ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*, (2) :168, August 2005.
- [Gra] Y. Le Grand. volume 1 & 3. Masson&Cie.
- [LCTS05] P. Ledda, A. Chalmers, T. Troscianko, and H. Seetzen. Evaluation of tone mapping operators using a high dynamic range display. *Proc. SIGGRAPH '05*, pages 249–258, 2005.
- [LRP97] G. Ward Larson, H. Rushmeier, and C. Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics*, pages 291–306, 1997.
- [MS45] P. Moon and D. Spencer. The visual effect of non-uniform surrounds. *Journal of the Optical Society of America*, 35(3) :233–248, 1945.
- [PFFG98] S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild, and D. P. Greenberg. Multiscale model of adaptation and spatial vision for realistic image display. *Proc. SIGGRAPH '98*, pages 287–298, 1998.
- [PTYG98] S. N. Pattanaik, J. Tumblin, H. Yee, and D. P. Greenberg. Time-dependent visual adaptation for fast realistic image display. *In Proceedings of ACM SIGGRAPH 2000*, pages 47–54, 1998.
- [RSSF98] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda. Photographic tone reproduction for digital images. *In Proceedings of ACM SIGGRAPH 2002*, pages 267–276, 1998.
- [VBC<sup>+</sup>02] F. Viénot, C. Boust, E. Da Costa, R. Brémond, and E. Dumont. Psychometric assessment of the look and feel of digital images. *Driving Simulator Conference 2002*, September 2002.
- [War94] G. Ward. A contrast-based scalefactor for luminance display. *Graphics Gems IV*, pages 415–421, 1994.

---

<sup>7</sup>Le simulateur de conduite de l'INRETS (Institut National de Recherche sur les Transports et leur Sécurité)