



Thierry BOREL est ingénieur de recherche pour la Société Technicolor, à Cesson-Sévigné, en Bretagne, et membre du Programme d'excellence 3D de cette société.

images – nous y reviendrons. Le premier succès de la 3D, *Chicken Little*, des studios Disney, sort en 2005, mais il faudra attendre *Avatar* de James Cameron, fin 2009, pour que la 3D déclenche un réel engouement. Le cinéma d'animation contribue notablement au déploiement rapide de la 3D : les trois quarts des films 3D des années 2000 en sont issus. Les images créées par ordinateur sont techniquement parfaites, notamment parce que les images gauches et droites ne présentent ni différence de colorimétrie ni distorsions géométriques. Autre avantage des films d'animation, aucun système compliqué d'acquisition à deux caméras n'est nécessaire.

Ce rapide historique met en exergue la notion clef de confort de visualisation. Une maîtrise insuffisante de certaines règles de production peut entraîner des fatigues oculaires, des maux de têtes ou des nausées. En outre, nous ne percevons pas tous un contenu stéréoscopique de la même façon. Cinq pour cent de la population ne voient pas le relief lorsqu'elle est exposée à ce type de contenu et 20 pour cent de la population le voient différemment : par exemple, certaines personnes ne perçoivent pas les sujets qui « sortent » de l'écran. Pour créer un contenu 3D de qualité, il est primordial de bien comprendre le fonctionnement du système visuel humain.

Lorsque le regard se fixe sur un objet, deux phénomènes se déroulent simultanément, chacun assuré par des muscles spécifiques : la mise au point (ou accommodation) et la convergence. La mise au point est une déformation du cristallin (qui joue le rôle de lentille convergente) ; les rayons lumineux sont alors focalisés sur la rétine, de sorte que l'image est nette. La convergence est un mouvement latéral des globes oculaires : ils se tournent vers l'objet, afin que leurs axes optiques se croisent à son niveau.

La lumière issue de l'objet est ensuite transformée en influx nerveux, qui sont transportés jusqu'au cortex visuel, dans la partie postérieure du cerveau. Les influx y sont analysés avec d'autres informations, telle l'intensité des contractions des muscles oculaires assurant la convergence et la mise au point, ce qui crée une représentation de l'espace environnant en couleur et en volume.

Avec l'accommodation, la convergence est un indice majeur, qualifié de binoculaire, qui permet au cerveau de percevoir le relief. D'autres indices, dits monoculaires, sont

également importants : les lignes de fuite, les occlusions (un objet partiellement occulté par un autre se situe derrière lui), les tailles relatives (plus un objet semble petit, plus il est éloigné), les textures (plus leurs détails sont visibles, plus ils sont proches)... À cette liste d'indices statiques s'ajoute un indice dynamique nommé parallaxe de mouvement et qu'on observe par exemple à travers la fenêtre d'un train : le mouvement apparent d'un objet éloigné de la fenêtre est plus lent que celui d'un objet proche. Grâce à ces indices monoculaires, on peut percevoir l'environnement en profondeur même lorsque l'un des deux yeux est déficient.

Des indices qui doivent être cohérents

L'appropriation des indices de profondeur monoculaires par les artistes a été progressive. Les peintres de la Renaissance ont été les premiers à maîtriser les lignes de fuite. Après l'invention de la photographie, on a ajouté un indice de profondeur artificiel : le flou artistique. En rendant flou l'arrière-plan ou l'avant-plan tout en laissant net le personnage principal, on guide l'attention du spectateur vers ce dernier et on accentue les différences de distance perçues entre les plans de l'image, ce qui augmente l'impression de profondeur. Tous ces indices ont été repris par le cinéma.

Pour restituer le relief sur un écran, il ne suffit donc pas de présenter une image différente à chacun des yeux : on doit aussi s'assurer que les indices de profondeur monoculaire et binoculaire concordent. C'est toujours le cas dans le monde réel, mais pas en stéréoscopie : si un objet jaillit de l'écran, par exemple, mais qu'il est masqué par le bord de cet écran (de sorte que l'indice monoculaire d'occlusion indique qu'il est derrière l'écran), cela crée un conflit psychovisuel gênant. Vérifier l'harmonie des différents indices est un travail important, qui doit être effectué avant la phase de tournage.

Entrons maintenant dans cette dernière phase. De même que les yeux reçoivent deux images différentes, on enregistre deux vues d'une même scène sous des angles légèrement distincts, à l'aide de deux caméras montées sur un même bâti (ou on crée ces deux vues dans le cas d'un film d'animation). Elles sont écartées horizontalement l'une de l'autre d'une distance dite interaxiale. Plus elle est grande, plus l'effet 3D sera important.

✓ BIBLIOGRAPHIE

P. Routhier, *Paving the road to 3D TV viability*, *ScreenPlays*, 2011.
www.screenplays-digital.com/screenplays/201109?pg=21#pg21

C. Vienne *et al.*, *Eye movements during vergence effort in stereoscopic ocular pursuit task and their relations to visual fatigue and stereopsis*, Vision Science Society annual meeting, 2011.

L. Blondé *et al.*, *3D stereo rendering challenges and techniques*, Conférence on information science and system, 2010.

L. Lipton, *Foundations of the Stereoscopic Cinema*, Van Nostrand Reinhold Inc., 1982.

Le plan de convergence est un plan qui passe par l'intersection des axes optiques des deux caméras. Les objets situés devant, dans et derrière ce plan seront respectivement visualisés devant le plan de l'écran, dans le plan ou derrière.

La longueur focale F (la distance entre l'objectif et l'image d'un point situé « à l'infini », c'est-à-dire en pratique éloigné d'au moins quelques dizaines de mètres des caméras) doit être aussi identique que possible pour les deux objectifs. Les rayons lumineux issus d'un même point ne frappent pas les capteurs des deux caméras au même endroit, et la différence de position dépend de la longueur focale et de l'angle de convergence α (l'angle entre les axes optiques des caméras). Cette différence, exprimée en pixels, est égale à la distance qui sépare les images gauche et droite d'un point sur l'écran lors de la projection. Si elle est trop importante, les yeux devraient diverger pour fixer les deux images, ce qui est impossible (pour la plupart d'entre nous !) : cet effet, dit de divergence, peut entraîner une fatigue visuelle, voire

la perception d'un dédoublement du fond de l'image dans les cas extrêmes.

Dès lors, on ajuste la longueur focale et l'angle de convergence en fonction de la taille maximale de l'écran visé. En effet, cette taille détermine la distance entre les pixels pour une image de résolution donnée (ayant par exemple 2048 pixels par ligne, un format classique), et donc le nombre de pixels d'écart autorisé entre les images droite et gauche d'un point. C'est une des différences clés avec un tournage 2D, puisqu'on doit connaître à l'avance la dimension maximale de l'écran pour lequel le contenu est prévu (écran de parc d'attraction, de cinéma, de téléviseur...) et s'adapter en conséquence.

La transformation d'un film prévu pour le cinéma en une version télévisuelle ne pose pas de problème de divergence puisque la distance entre les pixels, et donc entre les images gauche et droite d'un point, est plus petite sur un téléviseur ; l'inconvénient est que l'effet 3D sera moins saisissant. L'inverse n'est pas vrai : un contenu prévu pour la télévision, telle la retransmission 3D d'une rencontre sportive, néces-

site un traitement de postproduction important pour être projeté au cinéma.

Des problèmes de conversion se posent aussi avec les produits nomades, tels les caméscopes, les téléphones et les tablettes numériques : ces produits commencent à être équipés de doubles caméras, permettant à l'utilisateur de produire son propre contenu 3D. Or un contenu agréable à regarder sur une tablette risque d'entraîner des fatigues visuelles sur un écran de 50 pouces. On fournit donc des outils d'aide au tournage, afin d'apprendre au public à s'adapter au type d'écran visé.

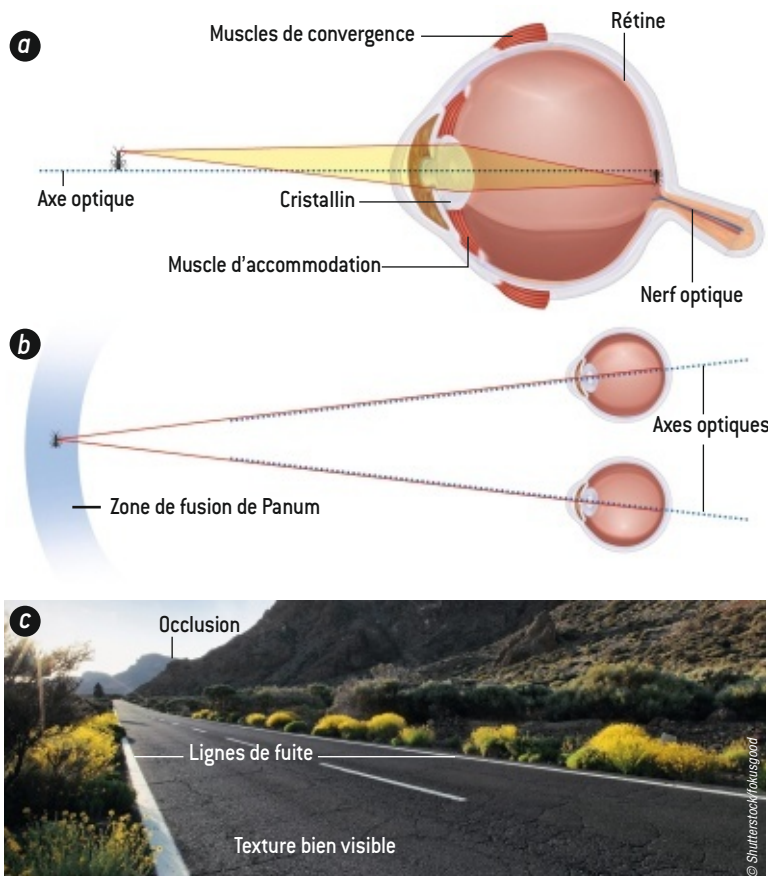
Lors du tournage et de la projection, on prête également attention aux proportions des objets. Pour un réglage de caméra donné, il existe une distance optimale à l'écran, pour laquelle les volumes sont respectés : à cette distance, l'angle moyen entre les deux yeux lorsque leurs axes optiques se croisent dans le plan de l'écran est égal à l'angle de convergence des caméras. Quand le spectateur s'éloigne de l'écran, il observe un allongement des objets et des visages ; on parle d'effet Pinocchio.

VOIR EN RELIEF : LES INDICES DE PROFONDEUR

Pour élaborer une représentation en relief, le cerveau analyse le degré de recouvrement des images légèrement différentes transmises par l'œil droit et l'œil gauche. Lorsqu'on fixe un objet, les yeux effectuent des réglages nommés accommodation et convergence. L'accommodation est une déformation du cristallin (en fonction de la distance de l'objet), permettant aux rayons lumineux issus de l'objet d'être focalisés sur la rétine (a). Sans cela, l'image est floue. Lors de la convergence, les deux globes oculaires se tournent vers l'objet, afin que leurs axes optiques se croisent dans le plan de l'objet (b).

Pour estimer l'éloignement de l'objet, le cerveau utilise les différences de localisation entre les images d'un même point sur les deux rétines, ainsi que les informations issues des muscles d'accommodation et de convergence. Autour de l'objet se trouve une zone dite de fusion de Panum : pour les éléments situés dans cette zone, le cerveau fusionne les images issues des deux yeux ; en dehors, le spectateur voit double. On parle de diplopie. Ainsi, quand on regarde au loin et qu'on rapproche un doigt de son visage, il se dédouble, car il sort de la zone de fusion de Panum.

Pour apprécier la profondeur d'une scène, le cerveau se fonde aussi sur les indices dits monoculaires (c), telles les lignes de fuite, les occlusions (un élément partiellement masqué par un autre, telle la montagne du fond, est situé derrière lui), les tailles relatives (les massifs de fleurs les plus éloignés apparaissent les plus petits) et les textures (plus elles sont visibles, plus l'objet est proche).



Bruno Bourgeois