

JACQUES AUMONT

# L'IMAGE

Peinture, photographie, cinéma :  
des origines à l'ère numérique

4<sup>e</sup> édition enrichie

**ARMAND COLIN**

## Du même auteur chez le même éditeur

*L'Analyse des films*, avec Michel Marie, 4<sup>e</sup> éd., 2020.

*Introduction à la couleur*, 2020.

*L'Interprétation des films*, 2017.

*Esthétique du film*, avec Alain Bergala, Michel Marie et Marc Vernet, 3<sup>e</sup> éd., 2016.

*Dictionnaire théorique et critique du cinéma*, avec Michel Marie, 3<sup>e</sup> éd., 2016.

*Les Théories des cinéastes*, 2<sup>e</sup> éd., 2011.

*Le Cinéma et la mise en scène*, 2<sup>e</sup> éd., 2010.

Conseiller éditorial : Michel Marie

Illustration de couverture : *Videodrome*, David Cronenberg, 1983

© Universal / Photo : Mary Evans / Aurimages

Direction artistique : Élisabeth Hébert

Composition : Nord Compo

Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.

Le Code de la propriété intellectuelle du 1<sup>er</sup> juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements

d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour

les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée.

Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du

Centre français d'exploitation du droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).



© Armand Colin, 2005, 2011, 2020

© Nathan, 1990 pour la 1<sup>re</sup> édition

Armand Colin est une marque de

Dunod Éditeur, 11 rue Paul-Bert, 92240 Malakoff

[www.armand-colin.com](http://www.armand-colin.com)

ISBN : 978-2-200-62821-5

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2<sup>o</sup> et 3<sup>o</sup> a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

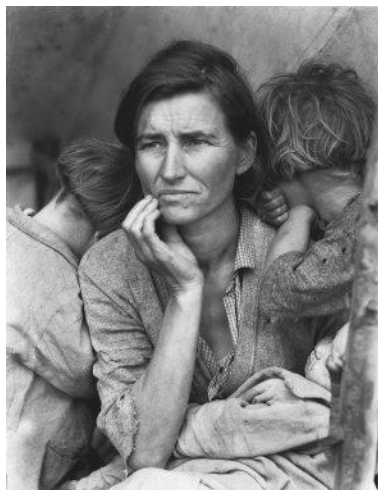
Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

# Introduction à la quatrième édition

Un jour de mars 1936, la photographe Dorothea Lange revient d'une mission dans le centre des États-Unis. Pour le compte de la Farm Security Administration rooseveltienne, elle vient de finir un reportage sur la misère des travailleurs des campagnes, ruinés par la grande dépression économique et par la désastreuse surexploitation des sols. Au volant de sa voiture, elle rentre chez elle, à New York, quand au bord de la route elle aperçoit un campement d'ouvriers agricoles condamnés au chômage. Elle fait demi-tour, pénètre sur le terrain, et comme poussée par une force inconnue, s'approche d'une tente qui abrite une jeune mère et ses quatre enfants, faméliques et désespérés. Sans un mot, en moins de dix minutes, elle prend cinq clichés, et repart. La dernière de ces photos, celle qu'elle prend de près, allait devenir, sous le titre *Migrant Mother*, l'une des plus célèbres photos du monde, reproduite encore et encore dans les journaux (où elle remplit sa mission de témoignage), puis dans des expositions (dont celle de Steichen, *The Family of Man*, en 1953), et se trouve aujourd'hui à la Library of Congress.

Cette histoire vraie pourrait être un apologue. Une image, une simple image, si les circonstances y prêtent, peut devenir un moyen important, efficace, puissant pour témoigner, pour montrer le monde tel qu'il va ou ne va pas. Souvent reproduite, elle peut devenir célèbre, tellement qu'à la longue on ne la voit presque plus, elle s'efface sous des traits trop connus (l'effet *Joconde*). Munie d'un titre un peu vague, générique, elle symbolise bien davantage que

la réalité contingente qui lui a donné son prétexte : elle peut devenir une icône. Cependant, si on prend la peine de la regarder, on verra que ni sa célébrité ni sa force ne sont de pure convention : cette image d'une jeune femme de trente-deux ans qui a l'air d'en avoir cent, de deux jeunes enfants cachant leur visage (de honte ?) et d'un bébé au visage difforme et maculé, cette pose de l'avant-bras dont on ne sait si elle trahit la perplexité, la gêne, la douleur – tout cela est immédiatement expressif, nous n'échappons pas à un mouvement de sympathie, d'émotion devant ce tableau de la dérélition.



Dorothea Lange, *Migrant Mother, Nipomo, Californie* (1936).

L'image, dans notre culture, c'est tout cela : une reproduction d'un aspect du monde, captée par la sensibilité plus ou moins individuelle, plus ou moins générique et typique d'un médiateur (artiste ou autre), qui nous communique des informations et des sentiments, et en suscite encore d'autres en nous – et c'est aussi un objet nouveau ajouté au monde, un objet qui vit sa vie, devient célèbre ou reste ignoré. L'image est un sosie du monde, mais un sosie déformé, si réaliste soit-il, et qui a sur nous l'attrait et l'efficacité magiques du double, en même temps que le charme propre d'un objet purement visuel.

En intitulant ce livre *L'image*, j'ai eu l'ambition de rendre compte de tout cela – la ressemblance, l'expressivité, l'émotion, la reconnaissance du

visible, la mise en forme, bref, un domaine vaste et divers de l'activité et de l'expérience humaines. Mais je me suis délibérément limité à l'image que l'on fabrique et que l'on regarde, c'est-à-dire à cet objet ajouté au monde par l'activité humaine, et qui sert à le représenter. Le mot « image », dans notre langue, désigne encore bien d'autres réalités, de l'image mentale à l'image onirique, de l'image poétique à l'imagerie symbolique. Tous ces sens, importants voire essentiels, sont hors des limites de ce petit livre, dont le domaine propre se restreint, par principe, aux images qui se réfèrent objectivement au monde (même si c'est parfois indirectement, comme la peinture « abstraite »). Je suis conscient des inconvénients de ce choix, qui d'ailleurs ne pourra être absolument tenu (il est presque impossible de ne jamais évoquer, à propos de la simple perception d'une image, l'« image mentale » ou l'imagination). Mais il m'a semblé le plus raisonnable, ce panorama étant déjà amené à aborder énormément de problèmes divers, souvent trop vite, et ne pouvant embrasser un domaine encore plus vaste.

La première édition de ce livre, fondée sur le même principe, était organisée en cinq grandes parties, qui correspondaient à un découpage moins disciplinaire que thématique. Trente ans plus tard, ce découpage me semble trop marqué par son époque pour être reconduit tel quel. Pour ne prendre qu'un exemple, je consacrais en 1990 une grande place à des théorisations du dispositif, appartenant au vocabulaire et aux concepts de la sémiologie d'inspiration psychanalytique des années 1970, et qui aujourd'hui ont tout à gagner à être mises davantage en perspective dans une considération, plus actuelle, du médium de l'image ou de son lieu. De même, j'avais donné à la physiologie de la perception visuelle une place qui à présent me semble excessive, au vu de ce qu'elle apporte réellement à une compréhension de ce qu'est l'image. Plus localement, j'ai aussi relativisé l'importance de telles problématiques (« technique et idéologie », « différences sexuelles et scotophilie », etc.), que je me suis efforcé de présenter sous un éclairage moins daté. Au total, j'ai ainsi préféré, pour cette nouvelle édition, adopter un plan plus parcellisé, autour de séries de questions dont l'enchaînement n'a rien de théorique.

Par ailleurs, une question se posait dans cette réécriture : quelle place accorder à deux considérations qui ont occupé le devant de la scène depuis 1990, celle de la « révolution numérique », et celle de la « médiologie » ?

L'image est-elle encore telle que nous la connaissons jusqu'à la photographie et au cinéma, c'est-à-dire une reproduction analogique du monde – ou bien l'invention de l'image numérique, l'apparition de procédés immédiatement vulgarisés d'intervention ou de retouche, ont-elles amené à une tout autre conception de ce qu'est une image ? Si beaucoup de spécialistes de l'image numérique et de l'image de synthèse plaident pour que soit reconnue à leur objet d'étude une spécificité absolue, j'ai toujours été, pour ma part, réticent devant cette idée : si je vois une photographie prise avec un appareil numérique, je vois une photographie, largement semblable à ce que je pouvais voir avant l'invention du numérique. Certes il y a des différences : je la vois plutôt sur un écran que sur papier ; je la jette sans états d'âme, parce que la produire ne coûte rien ; au prix d'une compétence minimale, je peux aussi la modifier. Mais ces différences ne me semblent pas suffisantes pour qu'on puisse parler d'une autre espèce, voire, comme c'est le cas ici ou là abusivement, d'une autre nature d'image. On ne trouvera donc pas, dans ce livre, de développement séparé consacré à l'image numérique – mais, bien entendu, je tiens compte partout de l'existence de cette importante nouveauté. Quant à la médiologie, malgré les aperçus brillants que l'on peut trouver, occasionnellement, dans tel ou tel ouvrage visant à la promouvoir (au premier chef ceux de son fondateur), elle ne m'a pas semblé avoir acquis incontestablement le statut d'une discipline autonome, et je me suis donc contenté d'en intégrer également les suggestions là où elles pouvaient être pertinentes.

Au total, je ne puis que réitérer l'avertissement de l'édition de 1990 : ce livre n'est qu'une introduction à des approches plus spécialisées, et c'est pourquoi il vise avant tout à signaler la multiplicité même de ces approches (y compris au sein d'une même discipline). Je ne suis d'ailleurs nullement spécialiste de toutes les questions que j'évoque, toujours trop succinctement ; j'espère seulement en avoir montré l'intérêt et permis qu'on en prolonge l'étude. En écrivant cette version comme la précédente, j'ai pensé avant tout aux étudiants que j'ai rencontrés pendant quarante ans à l'université ; c'est en sachant combien le savoir théorique qui leur est dispensé est parcellisé que j'ai tenté, non de leur proposer la clef définitive à toutes les questions, mais de leur permettre de situer ce qu'ils savent de l'image et des images dans une réflexion plus large. Bien entendu, je crois que cela peut intéresser un public bien plus divers – comme l'a démontré le succès de la première version.

Un dernier avertissement : il est évident qu'un tel livre, qui est plutôt un *compendium* qu'un essai suivi, peut se lire dans l'ordre qu'on voudra. Les chapitres et même les sous-chapitres ont été voulus largement indépendants – quitte à parfois aborder le même problème, de points de vue pas trop redondants je l'espère, en divers endroits du livre. Pour faciliter cette lecture et la circulation qu'elle implique, j'ai multiplié les références internes. J'ai aussi, à dessein, multiplié les références bibliographiques (sous forme abrégée, renvoyant à la bibliographie finale), en ne commentant qu'à peine les ouvrages, même importants, que je cite – à charge pour le lecteur de prolonger sa lecture en bibliothèque (ou, s'il y parvient, sur son écran).

Comme il y a trente ans, je dédie ce livre à Christian Metz (1931-1993), qui fut mon maître et mon ami, et le demeure.





# La perception de l'image

L'image telle qu'on l'entend ici est un objet visible et visuel. Nous commencerons, empiriquement, par nous demander comment elle est vue, et pour cela, par rappeler quelques données de base sur la perception visuelle.

## 1. LE PROCESSUS DE PERCEPTION

### 1.1 Œil et système visuel

L'expérience quotidienne, le langage courant, nous disent que nous voyons avec les yeux. Ce n'est pas faux, mais il faut ajouter qu'ils ne sont qu'un des instruments de la perception, et pas le plus complexe. La vision est un processus, qui met en jeu plusieurs organes spécialisés, et résulte de trois types d'opérations distinctes et successives : optiques, chimiques et nerveuses.

#### *Des opérations optiques*

L'œil est un globe quasi sphérique, couvert d'une couche en partie opaque, en partie transparente. C'est cette dernière, la cornée, qui assure le passage et la convergence des rayons lumineux. Derrière elle se trouve l'iris, muscle sphincter commandé de façon réflexe, percé en son centre d'une ouverture, la pupille. Celle-ci s'ouvre (de 2 à 8 mm environ) pour laisser pénétrer plus de lumière lorsque cette dernière est peu intense ; elle se ferme dans le cas contraire. La diminution de la pupille modifie la

perception : plus elle est fermée, plus la profondeur de champ est importante (c'est pourquoi on voit plus net lorsqu'il y a beaucoup de lumière).

La lumière qui a traversé la pupille doit encore traverser le cristallin. Celui-ci est, en termes optiques, une lentille biconvexe à convergence variable ; c'est cette variabilité qu'on appelle l'accommodation. Accommoder, c'est rendre le cristallin plus ou moins bombé, en fonction de la distance de la source de lumière : pour maintenir l'image nette sur le fond de l'œil, il faut faire d'autant plus converger les rayons que la source lumineuse est rapprochée. Comme l'ouverture de la pupille, c'est un processus réflexe, d'ailleurs assez lent, puisqu'il faut près d'une seconde pour passer de l'accommodation la plus proche à la plus lointaine.

On a très souvent comparé l'œil à une *camera obscura* (p. 151) ou un appareil photo en miniature ; cela est exact, à condition de bien voir que **cette comparaison ne s'applique qu'à la partie optique du traitement de la lumière**. Dans la *camera obscura* (boîte fermée seulement munie d'une petite ouverture), les rayons lumineux pénètrent, mais de manière diffuse, et l'image est pâle ; si on augmente la taille de l'ouverture pour laisser entrer plus de lumière, l'image devient floue. C'est pour pallier ce défaut que l'on a, dès le XVI<sup>e</sup> siècle, inventé les lentilles convergentes : des morceaux de verre taillés pour collecter la lumière et la concentrer en un point unique (p. 225). C'est ce principe de la capture d'un grand nombre de rayons sur une surface et de leur concentration en un point qu'utilise l'objectif photographique, et c'est ce même principe qui est à l'œuvre dans l'œil.

### *Des transformations chimiques*

Le fond de l'œil est tapissé d'une membrane, la rétine, dans laquelle se trouvent des organes récepteurs, en très grand nombre et de deux types : les bâtonnets (environ 120 millions) et les cônes (environ 7 millions) ; ces derniers sont surtout présents aux alentours de la fovéa, sorte de petit creux, à peu près sur l'axe du cristallin, particulièrement riche en récepteurs. Bâtonnets et cônes comportent chacun des millions de molécules d'une substance qui absorbe des quanta lumineux, et se décompose par réaction chimique. Une fois cette décomposition opérée, la substance en question ne peut plus rien absorber ; mais si l'on cesse de lui envoyer de la lumière, la réaction s'inverse et elle se recompose (il faut rester environ

trois quarts d'heure dans le noir pour que toutes les molécules de la rétine se recomposent, mais la moitié sont déjà recomposées au bout de cinq minutes) ; elle peut alors fonctionner à nouveau.

La rétine est donc avant tout un laboratoire de chimie miniature. Ce qu'on appelle « image rétinienne » n'est que la projection obtenue sur le fond de l'œil grâce au système cornée + pupille + cristallin, et cette image, qui est de nature optique, est traitée par le système chimique rétinien, qui la transforme en une information d'une nature totalement différente. Il est essentiel de bien comprendre que, contrairement à ce que peut suggérer le langage, **l'image rétinienne n'est pas une image** au sens où nous en parlons dans ce livre : elle n'est pas faite pour être vue, et d'ailleurs pas visible.

### *Des transformations nerveuses*

Chaque récepteur rétinien est relié à une cellule nerveuse par un relais (une synapse) ; chacune de ces cellules est, par d'autres synapses, reliée à son tour à des cellules qui constituent les fibres du nerf optique ; les communications entre ces cellules sont complexes, elles les constituent en réseaux. Le nerf optique sort de l'œil et aboutit à une région latérale du cerveau, d'où de nouvelles connexions nerveuses repartent vers l'arrière du cerveau, pour arriver au cortex strié. Schématiquement, on pourrait dire que ce réseau extrêmement dense représente un troisième et dernier stade, nerveux, du traitement de l'information ; ce traitement repose d'ailleurs lui aussi sur la chimie, les connexions nerveuses se faisant par échange de substances chimiques autour de la synapse. En règle générale, il n'y a pas correspondance de point à point, mais multiplication des correspondances transversales : le système visuel ne cesse de traiter l'information qu'il transmet.

Cette partie du système perceptif est la plus importante, mais aussi la moins bien connue, puisqu'on n'a commencé à avoir des idées un peu précises sur sa structure et son fonctionnement qu'au xx<sup>e</sup> siècle. En particulier, on ne sait toujours pas exactement comment l'information passe du stade chimique au stade nerveux, la nature même du signal nerveux, qui n'est que métaphoriquement comparable à un signal électrique, n'étant elle-même pas absolument claire. L'important est de retenir que, si l'œil ressemble jusqu'à un certain point à un appareil photographique, si la rétine est comparable à une espèce de plaque sensible, **l'essentiel de la**

**perception visuelle a lieu ensuite**, à travers un processus de traitement de l'information qui, comme tous les processus cérébraux, est plus proche de modèles informatiques ou cybernétiques que de modèles mécaniques ou optiques (Frisby, 1979).

## 1.2 Les éléments de la perception : que perçoit-on ?

La perception visuelle est le traitement, par étapes successives, d'une information qui nous parvient par l'intermédiaire de la lumière entrant dans nos yeux. Comme toute information, celle-ci est codée, ce qui signifie que notre système visuel est capable de repérer et d'interpréter certaines régularités dans les phénomènes lumineux qui atteignent nos yeux. Pour l'essentiel, ces régularités concernent trois caractères de la lumière : intensité, longueur d'onde, distribution dans l'espace (nous envisagerons plus loin la distribution dans le temps).

### *L'intensité de la lumière : la luminosité*

Ce que nous percevons comme la luminosité plus ou moins grande d'un objet correspond à une estimation, infléchiée par des facteurs psychologiques, de la quantité réelle de lumière émise ou réfléchiée par cet objet. Pour l'essentiel, l'œil réagit aux flux lumineux ; un flux même très faible, correspondant à une dizaine de photons seulement, peut suffire à faire enregistrer une sensation de lumière. Lorsque ce flux augmente, le nombre de cellules rétiniennes atteintes devient plus grand, les réactions se produisent en plus grand nombre, le signal nerveux devient plus intense. On distingue classiquement deux types d'objets lumineux, correspondant à deux types de vision, fondés sur la prédominance de l'un ou l'autre type de cellules rétiniennes :

a) La **vision photopique** : c'est le mode le plus courant, correspondant à la gamme des objets normalement éclairés par une lumière diurne. Il met en jeu surtout les cônes ; or ceux-ci sont responsables de la perception des couleurs, donc la vision photopique est *chromatique*. Les cônes sont surtout denses vers la fovéa, c'est donc aussi un mode de vision qui mobilise principalement la zone centrale de la rétine. Pour cette dernière raison, et parce qu'en outre la pupille peut alors être davantage fermée, vu la quantité de lumière, c'est une vision caractérisée par son acuité.

b) La **vision scotopique** : c'est la vision à faible lumière. Elle a les caractères opposés à la précédente : prédominance des bâtonnets, perception achromatique, de faible acuité, et concernant, surtout lorsqu'il fait très sombre, la périphérie de la rétine.

La vision des images, étant éminemment volontaire, met principalement en jeu la zone fovéale ; en outre, cette vision est presque toujours une vision photopique. Le cinéma, la vidéo, l'ordinateur, que l'on regarde dans le noir, en participent également, l'écran ayant une luminance élevée ; la vision de ces écrans est donc une vision diurne, même si elle a lieu dans un environnement sombre. Les cas de vision scotopique sont très rares en ce qui concerne les images, et toujours voulus, par exemple dans certaines installations artistiques.

### *La longueur d'onde de la lumière : perception de la couleur*

De même que le sentiment de luminosité provient des réactions du système visuel à la luminance des objets, le sentiment de couleur provient de ses réactions à la longueur d'onde des lumières émises ou réfléchies par ces objets. Contrairement à ce que suggère le langage, la couleur n'est donc pas exactement « sur les objets » ; elle est le résultat de notre perception de certaines qualités de surface des objets, et de la lumière qui les éclaire. La lumière blanche (notamment celle qui provient du soleil, première source lumineuse pour l'homme) est en fait un « mélange » de lumières, contenant toutes les longueurs d'onde du spectre visible, ce que l'on peut aisément mettre en évidence en la décomposant à l'aide d'un prisme, ou en observant un arc-en-ciel. La lumière qui nous parvient des objets est réfléchi par eux ; or, la plupart des surfaces absorbent certaines longueurs d'onde, et ne réfléchissent que les autres (celles qui absorbent toutes les longueurs d'onde paraissent grises).

L'expérience de la couleur a donné lieu à de nombreuses observations empiriques, depuis au moins le XVII<sup>e</sup>. On a ainsi de longue daté élaboré un système de description des couleurs selon trois paramètres :

1°, la longueur d'onde, qui définit la **teinte** (bleu, jaune, rouge, cyan, magenta, orange...) ;

2°, la **saturation**, c'est-à-dire la « pureté » (le rose est du rouge « moins saturé », i. e., auquel on a ajouté du blanc) ; les couleurs du spectre solaire ont une saturation maximale ;

3°, la **luminosité**, liée à la luminance : plus celle-ci est élevée, plus la couleur paraîtra lumineuse et proche du blanc ; le même rouge, également saturé, pourra être plus lumineux ou plus sombre.

Michael Snow a donné ce titre (*Hue Chroma Tint* = Teinte Saturation Luminosité) à une de ses œuvres (une installation de 2005).

Depuis Newton, on sait que les couleurs peuvent se *composer*. On distingue deux sortes de mélanges :

– **additifs**, ou mélanges de lumières. C'est le cas de la vidéo, où de très petits faisceaux de lumières en couleurs « primaires » (rouge, vert, bleu) se mélangent pratiquement sur la rétine, parce qu'ils sont trop étroits pour être perçus distinctement ; c'est sur un principe analogue que reposait le procédé autochrome de photographie en couleurs, autour de 1900, et aussi, en peinture, la tentative pointilliste, voulant rendre la couleur par une juxtaposition de très petites taches de couleur, fondues par l'œil en une seule couleur par addition ; c'est aussi sur le même principe que repose l'image numérique.

– **soustractifs**, ou mélanges de pigments. Comme chaque pigment ajouté absorbe de nouvelles longueurs d'onde, il s'agit bien d'une soustraction ; c'est le cas usuel de la peinture, ainsi que de la plupart des procédés de photographie et de cinéma en couleurs utilisant la pellicule.

Le principe général du mélange est le même dans les deux cas : deux couleurs mélangées produisent une troisième couleur ; avec trois couleurs, dites **primaires**, on peut obtenir n'importe quelle autre couleur, en dosant en conséquence les trois primaires.

La perception de la couleur est due à l'activité de trois variétés de cônes rétiniens, chacun étant sensible à une longueur d'onde différente (correspondant respectivement à un bleu-violet, un vert-bleu et un vert-jaune). Le codage de la couleur dans les étages terminaux du système visuel est très complexe ; certains groupements de cellules, depuis la rétine jusqu'au cortex, sont spécialisés dans la perception de la couleur, et celle-ci est donc l'une des dimensions essentielles de notre monde visuel. Cette dimension est absente de nombreuses images, qui sont achromatiques (« en noir et blanc ») : ce sont des images qui ne représentent pas les couleurs, mais seulement les luminosités, et qui comportent donc toute une gamme de gris. Le type le plus fréquent en a été l'image photographique, mais bien avant l'invention de la photo, des images « en noir et blanc »

circulaient dans notre société, avec notamment les gravures. C'est là un exemple majeur du fait que l'image représente la réalité de façon conventionnelle, correspondant à ce qui est acceptable socialement ; aucun des premiers spectateurs du Cinématographe Lumière, par exemple, ne s'est jamais vraiment plaint que l'image fût grise, et peu l'ont remarqué (Gorki, 1896) ; plus largement, le « noir et blanc » a longtemps semblé être la « couleur » naturelle du cinéma (Cavell, 1979).

### *La distribution spatiale de la lumière : les bords visuels*

Notre œil est également équipé pour percevoir les limites spatiales des objets, leurs bords. La notion de « **bord visuel** » désigne la frontière entre deux surfaces de luminance ou de couleur différente – quelle que soit la cause de cette différence (éclaircissements différents, propriétés réfléchissantes différentes, etc.). Il y a un bord visuel entre deux surfaces dont l'une est derrière l'autre, par exemple, mais si le point de vue change, le bord ne sera plus au même endroit ; c'est ce qu'on a parfois appelé « effet-écran » (p. 131).

Jusque vers 1950-1955, on considérait la perception en termes d'isomorphies successives, de l'objet au cerveau via l'œil, la rétine étant conçue comme un simple relais qui se contentait de transmettre l'information point par point sans l'interpréter. On pense aujourd'hui que le codage est ininterrompu de la rétine au cortex. Un même récepteur rétinien fait partie de centaines de champs différents, qui se recoupent mutuellement, selon des logiques variables (sommatrices ou inhibitrices). Sans entrer dans le détail, je retiens que le système visuel est équipé d'instruments capables de reconnaître un bord visuel et son orientation, une fente, une ligne, un angle, un segment ; ces percepts sont des unités élémentaires de notre perception des objets et de l'espace. Ce mécanisme extrêmement complexe donne de bons résultats, puisque nous sommes capables de distinguer des bords visuels de très petite taille<sup>1</sup>.

---

1. La mesure de l'acuité visuelle, chez un opticien, utilise un type de stimulus déjà élaboré – les lettres de l'alphabet –, pour lequel les performances de l'œil ne sont pas excellentes. Dans cette situation, une acuité de 1 (ou 10/10), correspondant à une discrimination angulaire de 1°, est considérée comme bonne. Mais si on donne à percevoir, non plus des lettres, mais une ligne noire sur fond blanc, l'acuité peut aller jusqu'à 120 (correspondant à un pouvoir séparateur d'une demi-minute d'angle).

C'est probablement en raison de cette importance des bords visuels dans notre perception naturelle que s'est développée, dans toutes les civilisations et dès le paléolithique (p. 207), la technique du dessin au trait. Il n'est pas évident, en effet, qu'on puisse établir une relation entre une ligne dessinée (ou peinte) et la perception d'objets réels dans leur environnement (Field, 1950), et la précocité de cette convention laisse supposer qu'elle s'appuie sur des propriétés innées de l'appareil perceptif.

### *Le contraste : interaction entre la luminosité et les bords*

Notre système visuel est, en fait, équipé pour détecter non pas tant des luminances que des changements de luminance ; la luminosité (psychologique) d'une surface est presque entièrement déterminée par sa relation à l'environnement lumineux ; elle est fonction, notamment, de son arrière-fond. Deux objets sembleront avoir la même luminosité si leur luminance relative par rapport à leur fond est la même, quelles que soient les valeurs absolues de ces luminances ; inversement, un même objet, identiquement éclairé (donc émettant la même luminance) sera jugé plus lumineux devant un fond plus sombre.



**Le contraste dans l'image peut devenir un moyen expressif, brouillant parfois la perception (Ingmar Bergman, *Les Fraises sauvages*, 1957).**



En outre, notre système visuel est capable de conjuguer la perception de la luminosité et des bords visuels. Ainsi, un contraste n'est perçu entre deux surfaces que si elles sont perçues comme étant dans le même plan ; si elles sont vues comme étant à des distances différentes de l'œil, leurs luminosités seront plus difficiles à comparer. De même, lorsqu'un bord visuel est perçu comme dû à l'éclaircissement (passage d'une zone éclairée à une zone d'ombre sur la même surface), et non à un changement de surface, la différence de luminance de part et d'autre du bord est systématiquement négligée, et les deux zones sont jugées également lumineuses.

L'important est que les éléments de la perception – luminosité, bords, couleurs – ne sont jamais produits isolément, mais simultanément, et que la perception des uns affecte la perception des autres. En ce qui concerne les images, elles sont perçues comme n'importe quel autre objet, et tout ce que nous venons de dire s'applique. En particulier, elles donnent de nombreux indices de surface, et les bords visuels y apparaissent presque systématiquement comme séparant des surfaces coplanaires ; par conséquent, le contraste entre ces surfaces sera davantage perçu, dans le cas d'une image figurative, qu'il ne le serait devant une scène réelle semblable à celle qui est représentée. Les peintres du clair-obscur ont abondamment joué de cette particularité : le contraste entre ombre et lumière représentées sur le tableau, chez Rembrandt ou le Caravage, est bien plus fort que dans la réalité ; un tableau vivant aura du mal à imiter la *Ronde de nuit* (la tentative de Godard dans *Passion* [1982] est plus proche du tableau que de la vision réelle, puisqu'il s'agit encore d'une image plane, cinématographique en l'occurrence).

### 1.3 L'œil et le temps

La vision est un sens spatial. Mais les facteurs temporels l'affectent grandement, pour trois raisons principales : 1°, la plupart des stimuli visuels varient avec la durée, ou se produisent successivement ; 2°, nos yeux sont en mouvement constant, faisant varier l'information reçue par le cerveau ; 3°, la perception elle-même n'est pas un processus instantané ; certains stades en sont rapides, d'autres beaucoup plus lents (les récepteurs rétiniens réagissent en moins de 1 ms (milliseconde) lorsqu'ils sont reposés ; en revanche, il s'écoule au moins 50 à 150 ms entre la stimulation du récepteur et l'excitation du cortex).

## *Variation dans le temps des phénomènes lumineux*

Je mentionne seulement deux phénomènes parmi les plus importants :

1°, **l'adaptation**. L'œil, nous l'avons vu, a une très large plage de sensibilité à la luminance, mais à chaque instant, dans la vie réelle, la gamme des luminances qu'il doit percevoir excède rarement le facteur 100 (de 1 à 100 cd/m<sup>2</sup> dans une pièce éclairée, de 10 à 1 000 dehors, de 0,01 à 1 la nuit...). Lorsqu'il est confronté à une variation brutale de la luminance, l'œil cesse de voir pendant un certain temps. L'adaptation à la lumière est d'ailleurs beaucoup plus rapide que l'adaptation au noir : il est pénible de sortir d'une salle de cinéma pour se retrouver brusquement en plein soleil, mais on récupère sa capacité de vision en quelques secondes, alors qu'en entrant dans la salle, il faut 30 à 40 minutes avant de bien voir (en dehors de l'écran, généralement assez lumineux). Ces phénomènes d'adaptation, mal expliqués au plan théorique, sont empiriquement très connus, et ont suscité de nombreux « trucs » destinés à pallier les lenteurs de l'adaptation au noir. (En voici un : si on veut lire une carte routière la nuit, sans perdre son adaptation au noir, il suffit de la lire avec des lunettes à verres rouges ; le rouge stimulant surtout les cônes, les bâtonnets resteront adaptés.)

2°, **le pouvoir de séparation temporelle**. L'œil ne perçoit deux phénomènes lumineux comme non synchrones que s'ils sont assez largement distants dans le temps : il faut au moins 60 à 80 ms pour les séparer à coup sûr, et cette durée passe à 100 ms (1/10 s) s'il faut en outre distinguer lequel est survenu le premier. Cette durée peut sembler courte ; elle est en fait fort longue si on la compare à d'autres performances sensorielles (le système auditif a une résolution temporelle de quelques microsecondes). Dans le même ordre d'idées, l'œil n'est pas très rapide à dénombrer les stimuli lumineux (au-delà de 6 à 8 flashes par seconde, par exemple, il ne perçoit plus des événements distincts, mais un continuum, par un phénomène d'intégration [p. 20]).

## *Les mouvements oculaires*

Non seulement nos yeux bougent à peu près en permanence, mais notre tête et notre corps sont également mobiles ; la rétine est donc en mouvement incessant par rapport à l'environnement perçu. Cette mobilité

continue elle semblerait devoir être une source de « bruit » visuel, dont il faudrait pallier les effets ; or, non seulement il n'en est rien, mais la perception est **dépendante** de ces mouvements. En stabilisant artificiellement l'image rétinienne par des dispositifs compliqués, on a constaté que cette stabilisation entraîne progressivement une perte de la couleur et de la forme, une sorte de brouillard qui vient masquer l'image peu à peu. Le mouvement de la rétine **est indispensable à la perception**.

Les mouvements oculaires sont de plusieurs types :

- les mouvements saccadiques, très rapides (environ un dixième de seconde), brusques, volontaires (par exemple lorsqu'on revient à la ligne dans la lecture), ou involontaires (pour aller examiner un stimulus détecté à la périphérie de la rétine).

- les mouvements de poursuite, par lesquels on suit un objet en mouvement : plus réguliers, plus lents, difficiles à réaliser en l'absence de cible mouvante.

- les mouvements de compensation destinés à garder la fixation durant le mouvement de la tête ou du corps ; ils sont entièrement réflexes.

- la dérive, sans objet assignable, et qui témoigne de l'incapacité de l'œil à garder une fixation ; c'est un mouvement de vitesse modérée et de faible amplitude, sans cesse corrigé par des microsaccades qui ramènent rapidement l'œil sur la fixation.

### *Facteurs temporels de la perception*

Le temps est donc **inscrit dans notre perception**. La connaissance a progressé considérablement, sur ce point, avec la découverte en 1974 de l'existence de deux types de cellules du nerf optique, les unes spécialisées dans la réponse « lente » à des états de stimulation permanents, les autres, dans la réponse à des états transitoires.

1°, **la réponse lente** est l'ensemble des effets de **sommation** ou d'**intégration** temporelle. Le premier effet, la sommation, se produit au niveau des récepteurs, qui, dans certaines limites, distingueront mal entre une lumière faible et assez longue, et un flash très court et intense, si la quantité d'énergie totale est la même. Le second effet est celui qui se produit lorsque plusieurs flashes sont intégrés en une seule perception parce qu'ils se succèdent très vite.

L'effet le plus connu dans ce domaine est ce qu'on appelle la **persistance rétinienne**, qui consiste en une prolongation de l'activité des récepteurs quelque temps après la fin du stimulus. Cette durée de persistance est d'autant plus grande que l'œil est davantage adapté au noir, c'est-à-dire plus reposé : la persistance peut alors atteindre plusieurs secondes pour un flash intense ; lorsque l'œil n'est pas adapté au noir, elle sera plus brève, de l'ordre d'une fraction de seconde. Un exemple souvent donné est celui du « cercle de feu » (observé dès l'Antiquité) : si on fait tourner assez vite une torche à bout de bras, on verra un cercle lumineux, sans pouvoir distinguer les positions successives de la flamme (cet effet, toutefois, n'est dû qu'en partie à la persistance rétinienne – et **n'a rien à voir avec une quelconque perception de mouvement**).

2°, la **réponse rapide** est l'ensemble des effets de réponse à des stimuli variant rapidement. Parmi ces effets, deux sont spécialement intéressants pour nous, parce qu'ils concernent la perception des images mouvantes :

– Le **scintillement** : tout se passe comme si le système visuel avait du mal à suivre les variations périodiques d'une lumière, lorsque la fréquence en est supérieure à quelques cycles par seconde, mais reste assez basse. Il se produit alors une sensation de papillotement, qu'on appelle le scintillement. Lorsque la fréquence des apparitions de la lumière augmente, cet effet finit par disparaître, au-delà de la « fréquence critique », et on perçoit alors une lumière continue. La fréquence critique est de l'ordre de 10 Hz pour des lumières d'intensité moyenne ; elle peut atteindre 1 000 Hz avec des intensités élevées.

Le scintillement se produit au cinéma lorsque la vitesse de projection est trop faible ; c'est entre autres pour l'éliminer que la vitesse de projection (donc aussi la vitesse de prise de vues) n'a cessé d'augmenter, passant de 12 environ à 16, puis progressivement 24 images par seconde. Lorsque l'intensité des lampes de projection augmenta, la fréquence critique devint supérieure à 24 Hz, et le scintillement réapparut. Pour l'éliminer sans augmenter encore la vitesse de projection – ce qui eût entraîné de sérieux problèmes mécaniques –, on usa d'une astuce, consistant à « dédoubler », voire à tripler le volet tournant du projecteur, coupant ainsi le flux lumineux du projecteur deux ou trois fois à chaque photogramme. Chaque photogramme est alors projeté deux ou trois fois, avant que la pellicule avance au photogramme suivant.

On passe ainsi, avec 24 images différentes par seconde, à  $2 \times 24 = 48$ , ou à  $3 \times 24 = 72$  images projetées par seconde, donc au-dessus de la fréquence critique. L'effet est différent avec le cinéma numérique, dont l'image est produite de manière continue (selon le principe de la vidéo), mais là aussi il faut une fréquence minimale pour éviter des sensations désagréables à l'œil.

– Le **masquage visuel** : des stimuli lumineux survenant à peu de temps l'un de l'autre peuvent interagir, le second perturbant la perception du premier ; c'est ce qu'on appelle l'effet de masque. Cet effet réduit la sensibilité au premier stimulus : on perçoit moins de contraste, l'acuité visuelle est moindre. Ce phénomène, souvent exploré depuis 1975, semble directement lié à l'activité différentielle des cellules transitoires et permanentes.

Au cinéma, on peut masquer la perception du mouvement par un stimulus « vide » en insérant un photogramme blanc entre deux photogrammes (en revanche, un photogramme noir – qui équivaut à une absence de stimulus – n'aura pas cet effet). Il est d'ailleurs possible que le masquage d'un photogramme par le suivant soit une des conditions qui favorisent la perception du mouvement, en éliminant la persistance rétinienne, mais cela est loin d'être certain, deux photogrammes successifs étant en général trop semblables pour pouvoir vraiment se masquer. En revanche, il y a un phénomène de masquage (donc, rappelons-le, transitoire par essence) au moment de la **coupe** ; certains films en ont abondamment joué, en multipliant les coupes abruptes. Cet effet est très réduit en numérique, le passage d'une image élémentaire à la suivante ne se produisant pas de façon aussi tranchée.

## 2. DU VISIBLE AU VISUEL

### 2.1 Perception de l'espace

C'est volontairement que je ne dis pas ici « perception visuelle » de l'espace : en effet, le système visuel n'a pas, à proprement parler, d'organe spécialisé dans la perception des distances, et la perception de l'espace n'est, dans la vie courante, presque jamais uniquement visuelle. L'idée d'espace

est fondamentalement liée au corps et à son déplacement ; en particulier, la verticalité est une donnée immédiate de notre expérience, via la gravitation : nous voyons les objets tomber verticalement, mais nous sentons aussi la gravité passer par notre corps. Le concept d'espace est donc autant d'origine tactile et kinésique que visuelle.

### *L'étude de la perception de l'espace*

Contrairement aux phénomènes élémentaires dont nous avons parlé plus haut, la perception de l'espace, phénomène complexe, ne s'étudie que très imparfaitement dans les conditions aseptisées du laboratoire.

Par exemple, la convergence entre les deux yeux est, intrinsèquement, une source d'information sur la profondeur ; on peut mettre son rôle en évidence en demandant à un sujet d'estimer la distance d'un point lumineux dans le noir complet. Mais, dans les conditions habituelles de vision, ce critère de profondeur n'est pratiquement jamais utilisé, et n'offre d'ailleurs qu'une information très peu sûre par rapport à d'autres critères plus puissants. Doit-on, dès lors, inclure la convergence entre les deux yeux dans une liste des indices de profondeur ? Répondre à ce genre de question demande que l'on prenne parti, plus nettement qu'à propos des percepts élémentaires, sur le modèle d'analyse adopté (p. 33 s.), et la conception que l'on a du processus perceptif dans son ensemble : la perception de l'espace est un domaine d'étude d'emblée théorique (Hochberg, 1978).

### *La constance perceptive*

À échelle macroscopique, les propriétés physiques du monde ne dépendent pas de notre regard sur lui : le monde a, *grosso modo*, toujours la même apparence, et nous nous attendons à y retrouver d'un jour à l'autre un certain nombre d'éléments invariants. C'est la perception de ces aspects invariants du monde (taille des objets, formes, emplacement, orientations, propriétés des surfaces, etc.) qu'on désigne par la notion de **constance perceptive** : malgré la variété des perceptions, nous repérons des constantes.

Cette notion est à rapprocher d'une autre idée, voisine, que l'on pourrait appeler la stabilité perceptive. Notre perception se fait par échan-

tillonnage continu (alternance de mouvements de l'œil et de fixations brèves) ; or, nous n'avons conscience, ni de la multiplicité de ces « vues » successives, ni du flou durant les mouvements oculaires, mais interprétons au contraire notre perception comme celle d'une scène stable et continue. Dans le même ordre d'idées, on peut ajouter encore d'autres observations : ainsi, l'image optique formée sur la rétine est très floue et très peu colorée sur ses bords ; or, la perception d'une scène est toujours panoramique, chaque point en est susceptible d'être vu (et nous gardons à l'esprit cette possibilité en permanence). De même encore, nous voyons avec deux yeux, mais nous n'avons de la réalité, sauf exception, qu'une seule image – d'où la notion d'un « œil cyclopéen », sorte de combinaison idéale des deux yeux (p. 28).

Au total, il est clair que la vision de l'espace est une activité bien plus complexe que la simple « prise en photo » rétinienne. La constance et la stabilité perceptives ne peuvent s'expliquer si l'on n'admet pas que **la perception visuelle met en jeu, quasi automatiquement, un savoir sur la réalité visible.**

### *La géométrie spatiale*

À notre échelle, l'espace physique peut se décrire à l'aide d'un modèle simple et ancien, celui de la géométrie à trois axes de coordonnées deux à deux perpendiculaires (les coordonnées dites cartésiennes). Ce modèle dérive, par perfectionnements successifs, de la géométrie euclidienne, laquelle se caractérise entre autres par le fait qu'elle décrit l'espace comme possédant trois dimensions. Ces trois dimensions peuvent se concevoir intuitivement, en référence à notre corps et à sa position dans l'espace : la verticale est la direction de la gravité et de la station debout ; l'horizontale est celle de la ligne des épaules, parallèle à l'horizon visuel devant nous ; la troisième dimension est celle de la profondeur, correspondant à l'avancée du corps dans l'espace. (Bien entendu, cette compréhension intuitive ne saurait être confondue avec la géométrie elle-même.)

Or, le premier stade, optique, du traitement de la lumière dans le système visuel consiste à former une image à deux dimensions, sur le fond de l'œil. L'image rétinienne est une représentation, par projection, des luminances des diverses surfaces de la scène vue, et la relation entre un objet et son « image » rétinienne relève des lois de l'optique géométrique. Ce

caractère projectif de l'image optique formée dans l'œil est connu depuis l'Antiquité. Euclide notait, vers 300 av. J.-C., que l'image d'un objet au fond de l'œil est plus petite si cet objet est plus loin. Dès le XVI<sup>e</sup> siècle (époque où la *camera obscura* était très en faveur), l'analogie entre œil et chambre obscure a été souvent relevée. Il faut noter cependant que, en toute rigueur, cette correspondance entre objet et image est assez complexe, puisque la projection se fait sur un fond sphérique. L'analogie avec la *camera obscura*, dont la paroi arrière est plane, est donc approximative.

Au reste, j'y insiste, l'« **image** » **rétinienne n'est qu'un stade du traitement de l'information lumineuse, et nous ne la voyons jamais**. Que cette image soit plane ou courbe importe donc infiniment peu, contrairement à ce qu'on a parfois pensé. Pour ne donner qu'un exemple Panofsky (1924) suppose que, nos images rétiniennes étant formées sur l'intérieur d'un sphéroïde, des horizontales situées devant nous seront vues conformément à la géométrie sphérique, et apparaîtront donc légèrement bombées, semblant converger sur les côtés de notre champ visuel. C'est accorder bien trop d'importance et de « réalité » à ce qui se passe sur la membrane intérieure de l'œil, et l'expérience confirme que nous percevons les droites comme des droites, et non des courbes.

L'optique géométrique monoculaire suffit déjà à fournir plusieurs informations que notre système visuel interprète ensuite en termes d'espace : ce qu'on appelle des **indices de profondeur**. En effet, le problème de l'espace visuel est essentiellement celui de la perception de la profondeur, les deux autres dimensions étant perçues de façon plus directe et moins ambiguë. Voici les principaux indices de profondeur monoculaires (pour un œil immobile) :

**1. Perspective linéaire** : les rayons lumineux passant par le centre de la pupille donnent de la réalité une image qui en première approximation (très proche de la vérité pour la zone fovéale) peut se décrire comme une projection sur un plan à partir d'un point, ce qu'on appelle perspective centrale, ou perspective linéaire. Dans cette transformation géométrique, les lignes parallèles à l'axe de vision (perpendiculaires au plan de projection) donnent, dans l'image, une famille de lignes qui convergent en un point (intersection de l'axe optique et de la rétine) ; la largeur de segments de droites parallèles au plan de l'image décroît, sur l'image, avec leur éloignement dans la réalité ; les éléments les plus éloignés de l'œil donnent



une image plus proche de l'axe de vision, etc. Ces lois sont simples et bien connues. Elles permettent de comprendre que cette transformation optique apporte une grande quantité d'informations sur la profondeur de la scène vue : la diminution apparente de taille sera interprétée comme un éloignement ; le rapprochement par rapport à l'axe optique (c'est-à-dire par rapport à l'horizon), également, etc.

Mais cette information n'est pas dénuée d'ambiguïté. On ne peut pas passer d'une scène à trois dimensions à une image bidimensionnelle sans perdre de l'information, et il n'est pas possible de « remonter » de la projection d'un objet sur un plan à cet objet lui-même : la même figure projetée correspond, en fait, à une infinité d'objets potentiels ; de nombreuses images jouent sur cette ambiguïté, par exemple pour induire un trouble sur la taille des objets représentés. Si l'œil interprète correctement les projections rétinienne dans l'immense majorité des cas, c'est, comme nous l'avons déjà vu, qu'il ajoute aux informations fournies par la perspective quantité d'autres informations indépendantes de celles-ci, qui peuvent les corroborer ou les infirmer. Combinée aux gradients de texture (voir ci-dessous), la perspective permet le plus souvent de distinguer aisément un angle ou un coin d'un objet du bord d'une surface.

J'ajoute qu'il faut se garder de confondre cette perspective, modèle de ce qui se passe dans l'œil (et pour cela appelée autrefois *perspectiva naturalis*), avec la perspective géométrique appliquée en peinture et, par la suite, en photographie, laquelle résulte d'une convention représentative en partie arbitraire, et doit être produite artificiellement (d'où son nom de *perspectiva artificialis*) – même si l'une et l'autre de ces perspectives se décrivent selon le même modèle géométrique (p. 37 s., p. 219 s.).

**2. Gradient de texture** : une scène visuelle comporte des objets sur un fond ; or, les surfaces de ces objets ont une structure fine, plus ou moins régulière, ce qu'on appelle leur texture apparente. Un mur de brique aura une double texture : une texture grossière, correspondant aux jointures des briques, une texture très fine, celle des micro-aspérités de la brique. Une prairie vue d'assez loin a une texture fine, de même, un sol caillouteux. Naturellement, les objets faits de main d'homme ont en général des textures plus régulières (c'est le cas des tissus, précisément, dont le nom même évoque l'idée de texture). Les surfaces perçues étant habituellement inclinées par rapport à notre axe de vision, la projection