

SOCIÉTÉS & REPRÉSENTATIONS

numéro 55

L'œil numérique

Dossier dirigé par Laurence Danguy et Julien Schuh

ÉDITIONS DE LA SORBONNE



Laurence Danguy et Julien Schuh

L'œil numérique : vers une culture visuelle hybride

Il règne une certaine confusion lorsqu'on parle d'images numériques. Ce terme désigne en effet, chez ses détracteurs comme chez ses émules, des modes de production et des processus extrêmement variés, dont l'histoire, les enjeux et les conséquences sur la culture visuelle n'ont rien de commun. Cette introduction peut être vue comme une tentative de débrouiller l'écheveau des représentations enchevêtrées sous les vocables « images numériques », virtuelles, ou encore générées par ordinateur. Derrière les technologies concernées, des généalogies intellectuelles très différentes apparaissent, expliquant les impasses d'un discours unifié sur ces objets.

Comment parler des images numériques ? Les approches anthropologiques partent le plus souvent des images, allant jusqu'à leur donner une agentivité¹, sans prendre en compte (ou de manière superficielle) les processus humains et matériels qui seuls expliquent leur production, leur circulation, leur efficacité ; les études en culture visuelle, à rebours, analysent les discours autour des images (quitte à diluer la définition de leur objet en incluant les images verbales ou métaphoriques) mais tiennent tout autant à distance les conditions pragmatiques de leur existence. Le concept d'« œil » permet précisément d'articuler ces différents aspects : il peut subsumer l'ensemble des processus (cognitifs, matériels, technologiques, sociaux, industriels, mémoriels, institutionnels) qui font émerger les cultures visuelles. L'œil d'une époque, c'est une série de pratiques, de discours, d'objets, d'outils². Étudier « l'œil

1. Philippe Descola, *Les formes du visible*, Paris, Seuil, 2021, p. 23.

2. Voir W. J. T. Mitchell, *Iconologie : image, texte, idéologie*, traduit par Maxime Boidy et Stéphane Roth, Paris, Les Prairies Ordinaires, coll. « Penser/croiser », 2009 ; Maxime Boidy, *Les études visuelles*, Saint-Denis, Presses universitaires de Vincennes, 2017.

numérique », c'est tenter de comprendre la manière dont le numérique réordonne ces éléments et leurs interactions – non pas simplement en produisant de nouveaux formats d'image, mais en s'insérant profondément dans les mécanismes de production, de réception, de circulation, de sélection de notre culture visuelle.

Nous examinerons le sens à donner au « numérique » des images selon trois perspectives : celle de leur numérisation, la translation d'une image en chiffres, offrant de nouvelles potentialités ; celle de leur intégration dans un écosystème numérique, dont les mécanismes algorithmiques offrent des manières alternatives de contrôler le visible ; celle de l'émergence d'une culture visuelle hybride, où des images analysées et générées par des outils d'intelligence artificielle reconfigurent notre imagerie.

L'encodage des images

Le premier angle sous lequel envisager les images numériques est précisément celui de leur numérisation, c'est-à-dire des opérations de transformations des surfaces graphiques en code mathématique, offrant de nouveaux usages de l'image. La généalogie de la numérisation des images est intimement liée à la question de leur reproduction : d'une certaine manière, toutes les techniques de reproduction mécanique des images, depuis la xylographie jusqu'à sa captation numérique, sont des formes d'encodage. La gravure réduit la complexité des tableaux et des dessins selon des codes visuels (hachures, pointillés) ; la photographie, d'abord développée pour reproduire des œuvres d'art, opère selon des contraintes optiques et matérielles (focale, sensibilité de la pellicule, cadrage³). Mais la reproduction des images reste liée, dans ces techniques, à une matrice matérielle (bois gravé, pierre lithographique, négatif) et à des codes non systématisés. L'image numérique apparaît dès qu'un encodage permet de traduire en unités digitales la surface analogique d'une image, permettant son inscription (et sa transmission) par n'importe quel médium :

Une image numérique est un codage d'une réalité plane, la projection dans un plan d'une réalité tridimensionnelle, propriété de tout objet palpable. La notion est antérieure à l'ère de l'informatique. Il est depuis longtemps possible de transmettre une image de manière numérique. Il convient de décomposer l'image en

3. Voir Rémi Blachon, *La gravure sur bois au XIX^e siècle : l'âge du bois debout*, Paris, Éditions de l'Amateur, 2001 ; Stephen Bann, *Parallel Lines: Printmakers, Painters, and Photographers in Nineteenth-Century France*, New Haven/Londres, Yale University Press, 2001.

petits morceaux et d'attribuer un code couleur (une valeur numérique ou un caractère associé à la teinte) à chaque fragment. Ce code peut alors, en utilisant par exemple les conventions du morse, permettre la transmission de l'image (le morse est un code alphanumérique qui, à l'époque, était transmis de manière analogique). Lors de la réception il suffit de reconstituer l'image⁴.

Les premières entreprises de télétransmission des images, au XIX^e siècle, reposent sur une forme d'encodage rudimentaire. Dès 1858, le pantélégraphe de Caselli permettait de transmettre par fil des dessins, en les réduisant selon un code binaire (noir/blanc) grâce à l'utilisation d'un pendule qui parcourait rangée après rangée une plaque de cuivre sur laquelle on avait dessiné ou écrit au crayon gras non conducteur ; mais cet encodage ne pouvait être sauvegardé, et l'utilisation de cette machine nécessitait une transmission en direct, avec un autre pendule synchronisé sur le premier, qui redessinait l'image de départ sur une feuille imprégnée d'une substance réagissant au courant électrique⁵. Le bélinographe, en 1907, repose encore sur ce type de mécanismes, n'utilisant des cellules photoélectriques qu'en 1927, réduisant les images à des grilles de carrés noirs ou blancs⁶. L'histoire subséquente des évolutions de l'image électronique, où les frontières entre analogique et numérique sont parfois floues, passe de l'utilisation de bandes magnétiques aux cellules photoélectriques des appareils photo numériques⁷ ; comme le rappelle Claus Gunti, ce n'est que dans les années 1980 que se stabilisent les technologies de la photomosaïque standardisant l'usage du pixel⁸. L'aspect le plus important de ces techniques, indépendamment de leurs spécificités et de leur caractère purement numérique ou non, consiste dans ce qu'on peut appeler leur processus de virtualisation de l'image : non pas parce qu'à travers leur utilisation l'image cesse d'être matérielle (il y a toujours une trace matérielle, sous forme de traces électromagnétiques ou d'inscriptions textuelles à décoder), mais parce qu'elle cesse d'être visuelle au sens humain, c'est-à-dire immédiatement perceptible par nos sens, et peut faire l'objet d'une diffusion beaucoup plus large, la surface de

4. Nicole Vincent et Camille Kurtz, *Images et patrimoine numérisé : enjeux et nouveaux usages*, Lormont, Le Bord de l'Eau, 2019, p. 39-40.

5. Julia Zons, *Casellis Pantelegraph : Geschichte eines vergessenen Mediums*, Bielefeld, Transcript, 2015.

6. Myriam Chermette, « Transmettre les images à distance. Chronologie culturelle de la téléphotographie dans la presse française », *Études photographiques*, n° 29, 2012, p. 136-169. Voir également Jonathan Dentler, « Images câblées. La téléphotographie à l'ère de la mondialisation de la presse illustrée », *Transbordeur. Photographie histoire société*, n° 3, 2019, p. 14-25.

7. Peter Sachs Collopy, « La vidéo et les organes de la photographie électronique », *Transbordeur. Photographie histoire société*, n° 3, 2019, p. 26-35.

8. Claus Gunti, « Photographie », dans Raphaël Baroni et Claus Gunti (dir.), *Introduction à l'étude des cultures numériques : la transition numérique des médias*, Malakoff, Armand Colin, 2020, p. 87-103.

l'image n'existant plus qu'à travers son affichage par des dispositifs de décodage, rendant chacune de ses manifestations singulières. Comme le résume Hans Belting :

Face à un médium qui s'est lui-même désincarné, on assiste à la dissolution du lien physique entre l'image et son médium, alors qu'il fondait encore, jusqu'au tirage photographique, la règle de tout dispositif analogique. Dans le cas des images numériques, ce lien est en tout cas masqué par la dissociation entre appareil de production, la boîte noire de l'ordinateur, et dispositif de perception, le moniteur ou l'écran. Les images numériques sont stockées de manière invisible dans le système informatique. Quelle que soit leur apparence, elles renvoient à une matrice qui n'est plus une image. Leur caractère médial s'en trouve opacifié, à la fois élargi et discontinu, au sens que peut avoir le concept mathématique de valeur discrète⁹.

Cette condition nouvelle de l'image numérique concerne en réalité l'ensemble des objets culturels médiatisés par le numérique, introduisant une « rupture sémiotique » produite par la nécessité de passer par une interface impliquant un dispositif technique, une source d'énergie et une couche logicielle¹⁰. Mais, plutôt que de considérer ces modalités d'existence des images numériques comme une forme de dégradation d'une expérience originellement plus directe de la culture visuelle, il faut insister sur les affordances offertes par ces contraintes matérielles à travers deux aspects essentiels de l'image numérique, découlant de cette numérisation (au sens de réduction à des chiffres) : sa distribuabilité et sa manipulabilité.

La virtualisation des images, en permettant de les traduire en informations susceptibles d'être reproduites ou diffusées de manière très souple, constitue le stade ultime des recherches sur la reproduction des images et leur diffusion : la matrice matérielle, nécessaire auparavant, devient mathématique et informationnelle. L'image numérique permet de démultiplier les copies (parce que, grâce à la décorrélation entre le code et le support, il s'agit uniquement de recopier des données sur des mémoires électroniques, opération beaucoup plus simple que les techniques précédentes de reproduction photomécanique) et d'accéder à distance à ces informations pour les afficher sur un écran. C'est cette articulation entre virtualisation et conservation, permettant précisément une circulation inédite, qui fait événement, avec la distribution d'interfaces de consultation interconnectées, et génère une distribuabilité nouvelle.

9. Hans Belting, *Pour une anthropologie des images*, traduit par Jean Torrent, Paris, Gallimard, 2004, p. 54.

10. Emmanuel Souchier, Étienne Candel, Gustavo Gomez-Mejia, avec la collaboration de Valérie Jeanne-Perrier, *Le numérique comme écriture. Théories et méthodes d'analyse*, Malakoff, Armand Colin, 2019, chapitre 1.7.

L'autre caractéristique, celle de la manipulabilité de ces images, découle également de leur mathématisation : réduites à des chiffres, elles peuvent faire l'objet de traitements statistiques. Les logiciels de PAO type Photoshop et les applications de filtres sur les images constituent des outils d'exploitation de cette manipulabilité qui s'inscrivent dans une généalogie de retouche manuelle des images. Mais le développement de la vision par ordinateur a permis l'automatisation de la reconnaissance du contenu symbolique des images et de sa modification : pour la première fois, on peut déléguer des tâches de sélection, de classification, de censure, de transformation des images en se fondant sur des algorithmes et des modèles d'intelligence artificielle pré-entraînés. Cette mécanisation de la cognition a des conséquences extrêmes dans la production même des nouvelles images photographiques : nos téléphones portables, qui constituent désormais les premiers outils de production photographique, intègrent des modules de traitement automatique des images de plus en plus puissants, qui représentent une médiatisation d'un type nouveau dont les conséquences n'ont pas encore été véritablement pensées. L'appareil photo du Google Pixel ajoute ainsi une surcouche de traitements numériques à toute photographie : correction du flou, de la luminosité, des teintes de peau, imagerie à grande gamme dynamique (combinant des images prises simultanément à des valeurs d'exposition différentes), gomme « magique » permettant de supprimer ou de neutraliser des éléments du décor, reconnaissance des textes, classification des sujets et des visages pour une recherche ultérieure, etc. L'image photographique numérique se donne ici comme une empreinte alors qu'elle est le résultat de traitements algorithmiques qui modifient en profondeur la prise de vue.

Ces affordances ne sont cependant permises que par l'inscription de ces images numériques dans des infrastructures connectées, comprenant des réseaux, des dispositifs de consultation, des protocoles d'échange, des applications distribuées. Autrement dit, l'image numérique ne peut pas être comprise uniquement en termes techniques : c'est un complexe matériel, social, économique, politique.

Algorithmique : vers une culture visuelle post-humaine

Tenter de comprendre l'image numérique en tant qu'objet technique mène inexorablement à une exploration de l'écosystème qui permet sa production, sa conservation, sa diffusion. Si l'œil numérique peut désigner l'ensemble des processus de production, de transformation et de captation des images

digitales, il n'existe que dans le réseau d'acteurs, de protocoles, de dispositifs, d'infrastructures interconnectés qu'on désigne sous le terme « numérique » ou « internet » et qui organise ces processus – ce que Peter Szendy, après Andrew Ross, condense sous l'expression « écologie des images¹¹ ». Le numérique peut ainsi être considéré comme le nouveau lieu des images. L'œil numérique, dans ce contexte, représente les nouveaux mécanismes de traitement et de gestion de l'imagerie, c'est-à-dire de l'ensemble des images qui circulent dans une société donnée : c'est un outil de contrôle du visible.

Ces outils numériques s'inscrivent dans une tradition de régulation du visible et s'articulent avec des institutions et des protocoles existants. En réalité, c'est l'ensemble des dispositifs de traitement de la mémoire visuelle (dispositifs au sens foucauldien¹²) qui apparaît comme une vaste entreprise de computation des images¹³, c'est-à-dire de sélection, de transformation, de mise en circulation, de génération de la mémoire visuelle¹⁴. Ces opérations de conservation, de sélection, d'interprétation des images ont lieu de manière interactionnelle à travers une quantité de micro-transactions humaines ou automatisées (classification, hiérarchisation, transformation, destruction, reproduction) dans les institutions, les interactions sociales, économiques, politiques. Elles produisent de manière émergente des régularités et organisent la culture visuelle d'une époque. Les institutions (Beaux-Arts, INHA), les lieux de mise en scène de l'image (musées, cinémas, galeries, affichages, réseau de télévision), les formes médiatiques de production, de stockage (papier, celluloïd, négatifs), les outils de visionnage¹⁵ forment les rouages d'une vaste entreprise de configuration des regards mais aussi de production d'une culture des images où la part de l'œil humain est en partie décentrée. Diana Kamin a par exemple analysé la manière dont les conditions matérielles de conservation et de circulation des photographies dans les archives formaient une « iconomie », une base de données analogue dont les mécanismes donnent forme à une certaine culture visuelle¹⁶.

11. Peter Szendy, *Pour une écologie des images*, Paris, Éditions de Minuit, 2021, p. 34 ; il cite Andrew Ross, « The Ecology of Images », *The South Atlantic Quarterly*, vol. 91, n° 1, 1992, p. 215-238.

12. Voir Giorgio Agamben, *Qu'est-ce qu'un dispositif ?*, traduit par Martin Rueff, Paris, Éditions Payot & Rivages, 2007.

13. Sur la notion de computation, voir Edgar Morin, *La méthode*, Paris, Seuil, coll. « Opus », 2008, vol. 2, p. 1209.

14. Sur les conséquences du numérique sur les logiques mémorielles, voir Rémy Besson et Sébastien Ledoux (dir.), « Le numérique comme environnement mémoriel », *Mémoires en jeu*, n° 14, 2021.

15. Voir par exemple Bruno Latour, « Les "vues" de l'esprit ; une introduction à l'anthropologie des sciences et des techniques », *Réseaux*, vol. 5, n° 27, 1987, p. 79-96. DOI : 10.3406/reso.1987.1322.

16. Diana Kamin, « Cards, Cabinets, and Compression in Early Stock Photography », *Information & Culture*, vol. 56, n° 3, 2021, p. 1-23.

Le traitement numérique des images s'inscrit dans ce système même (il rajoute des rouages à la machine), mais plus essentiellement il opère un décentrement, une forme d'obfuscation¹⁷ qui rend le processus de computation encore plus invisible tout en étant d'une rapidité et d'une profondeur (d'une densité) bien plus importantes. D'analogique et diffuse, la computation des images devient numérique et ciblée, chaque image pouvant faire l'objet d'un traitement très complexe, selon des procédures distribuées grâce à l'ubiquité de la diffusion des codes à travers les dispositifs techniques. La puissance des outils de reproduction et de traitement des images mentionnés précédemment n'existe que parce que ces outils sont interconnectés et distribués à travers des milliards d'interfaces. Les logiciels de traitement automatique des photographies de nos smartphones sont inclus d'office dans leurs systèmes d'exploitation. Ils sont distribués massivement, mis à jour de manière globale : lorsque Google ou Apple modifient les algorithmes de traitement des photographies, c'est simultanément l'ensemble des appareils qui changent leur manière de produire des images. On oublie aussi souvent que l'image numérique n'est pas seulement l'encodage d'une image : elle comporte tout un ensemble de métadonnées qui contextualisent cet objet (date, localisation, appareil utilisé, format), métadonnées indexables qui en font un objet immédiatement disponible pour des recherches et des mises en série¹⁸.

De quelle nature est le contrôle exercé par cet « œil numérique » algorithmique ? Ce qui importe ici relève moins de la production ou de la transformation des images que de la régulation de leur circulation. Le développement d'outils de vision par ordinateur permet d'analyser le contenu des images et d'en réguler la diffusion en fonction de ces contenus : d'en interdire certains, d'en valoriser d'autres, de guider les flux attentionnels. Pour la première fois, les mécanismes de contrôle des contenus visuels reposent non pas sur des acteurs humains, mais sur des formes de cognition automatique produites par des machines. Les outils d'intelligence artificielle pourraient en effet plus pertinemment être appelés des outils de « cognition artificielle¹⁹ » : les algorithmes et les réseaux neuronaux au cœur des protocoles de censure ou de sélection des contenus médiatiques susceptibles d'être montrés aux utilisateurs de plateformes comme Youtube, Twitter, Instagram ou Netflix produisent de manière automatique, à une échelle incommensurable, des opérations

17. L'obfuscation désigne l'opération qui rend un programme informatique illisible aux humains tout en le maintenant fonctionnel.

18. Voir Lev Manovich, « "Metadating" the Image », dans Joke Brouwer et Arjen Mulder (dir.), *Making Art of Databases*, Rotterdam, V2 Pub./NAI Publishers, 2003, p. 12-27.

19. François Chollet, *Deep Learning with Python*, Shelter Island, Manning, 2021 [2017], p. 450.

cognitives originellement réservées aux humains. Ils utilisent la manipulabilité des images numériques (leur capacité à être analysées par des traitements statistiques) pour en contrôler la diffusion. Ces algorithmes participent désormais à la définition de nos cultures visuelles. Ils créent même des « cultures visuelles invisibles²⁰ » pour les humains, en utilisant des images qui ne nous sont plus destinées (vidéos de caméras de surveillance dans des entrepôts ou sur des autoroutes, panneaux de signalisation captés par les appareils des voitures autonomes de Google qui enrichissent Google Street View).

Les algorithmes qui traitent les images sur les réseaux sociaux ou les plateformes multimédiatiques (Netflix, Disney+) opèrent en analysant les images et les pratiques des abonnés pour contrôler de façon dynamique ce qui est visible. Ils fonctionnent de manière simultanément globale – ils peuvent être déployés partout à la fois et mis à jour en temps réel – et locale – ils produisent des effets différents selon les lieux de consommation des images, ou plus exactement, selon les niches ou les réseaux de consommateurs concernés, la localisation physique n'étant qu'un critère parmi d'autres, liés à des conditions légales ou politiques, dans une économie culturelle globalisée. Que ces algorithmes « voient » ou pas les images, au sens que l'on peut donner à ce terme en fonction de nos habitudes de perception, est relativement secondaire : ils participent directement à la forme de nos cultures visuelles.

L'un des enjeux essentiels de ces outils est de permettre la démultiplication de niches visuelles susceptibles d'être captées : l'histoire des médias visuels (en particulier la télévision, à travers la prolifération des chaînes câblées et désormais le streaming) montre que l'on est passé d'une tentative de proposer les mêmes images à tous (le « synopticon » de Thomas Mathiesen²¹) à celle d'offrir à tous les images qu'ils préfèrent (et qu'on leur fait préférer) en agrégeant de grandes tendances. Dans une forme d'automatisation des « techniques de l'observateur » analysées par Jonathan Crary au XIX^e siècle²², l'œil numérique peut aussi être compris comme une prothèse qui surajoute à nos sens une couche matérielle et algorithmique contrôlant le flux des images selon une logique de maximisation de la captation de l'attention visuelle, pour permettre sa revente²³.

20. Trevor Paglen, « Invisible Images (Your Pictures Are Looking at You) », *thenewinquiry.com*, 8 décembre 2016, cité par Peter Szendy, *Pour une écologie des images*, op. cit., p. 30.

21. Thomas Mathiesen, « The Viewer Society: Michel Foucault's "Panopticon" Revisited », *Theoretical Criminology*, vol. 1, n° 2, 1997, p. 215-234. DOI : 10.1177/1362480697001002003.

22. Jonathan Crary, *Techniques de l'observateur : vision et modernité au XIX^e siècle*, traduit par Frédéric Maurin, Bellevaux, Éditions Dehors, 2016.

23. Yves Citton, *L'économie de l'attention : nouvel horizon du capitalisme ?*, Paris, La Découverte, 2014.

L'œil numérique, c'est ainsi ce système qui nous voit regardant, qui observe notre regard et en tire des statistiques pour mieux l'attirer, à travers les outils de contrôle de la lecture (Amazon, par exemple, synthétise toutes les données issues de l'utilisation de ses *ebooks* et *ereaders*), les algorithmes des plateformes de streaming, le suivi des mouvements visuels intégrés dans les panneaux publicitaires numériques. Mais ces outils ne permettent pas seulement un contrôle des regardants, un façonnage de leurs habitudes de consommation visuelle : ils constituent également les rouages d'une forme de surveillance et de contrôle des regardés. Société du spectacle et société de surveillance fonctionnent depuis longtemps de concert : « L'opposition de Foucault entre le spectacle et la surveillance devient insoutenable : les deux modèles qu'il distingue se replient l'un sur l'autre. La production de l'observateur au XIX^e siècle coïncide avec de nouvelles procédures de discipline et de gestion²⁴. » Les outils numériques de traitement des images ont connu un développement accéléré dans les cadres militaires et policiers. La banalisation, en Chine, des caméras de surveillance et du contrôle automatique de la population fondé sur la reconnaissance des visages et des actions permet d'émettre des amendes, de faire baisser les scores citoyens, de surveiller l'attention des élèves dans des classes munies de caméras intégrées jusque dans les stylos des élèves. Comme l'analyse Adam Greenfield, ce système de surveillance globale des populations par la vision par ordinateur automatise le principe de contrôle visuel en place dans les grandes villes, où c'est l'entrecroisement des regards de tous qui constitue un système de régulation efficace des interactions selon Jane Jacobs²⁵.

C'est bien une culture visuelle au-delà du regard humain qui s'est mise en place, une imagerie reconfigurée par le numérique, où les mécanismes définissant ce qu'on peut voir, quand et comment, sont en partie délégués à des formes de cognition non humaines. On est ainsi passé des non-lieux (physiques, et marqués par l'anonymat) de Marc Augé²⁶ à des lieux (virtuels et physiques, de toute sorte) sous contrôle numérique.

24. Jonathan Crary, *Techniques de l'observateur*, op. cit., p. 169. Voir également Maxime Boidy, « *Visual Culture Studies* : les matérialismes du visible », dans Maxime Cervulle, Nelly Quemener et Florian Vörös (dir.), *Matérialismes, culture & communication*, tome 2, *Cultural studies, théories féministes et décoloniales*, Paris, Presses des Mines-Transvalor, 2016, p. 125-139.

25. Adam Greenfield, « China's Dystopian Tech Could Be Contagious », *The Atlantic*, 14 février 2018. Voir Jane Jacobs, *The Death and Life of Great American Cities*, New York, Vintage Books, 1961, p. 54.

26. Marc Augé, *Non-lieux – introduction à une anthropologie de la surmodernité*, Paris, Seuil, 1992.

Une culture visuelle hybride : les images comme espace latent de représentation

Enfin, c'est au niveau de la génération des images que l'on peut considérer les conséquences du numérique sur la culture visuelle, avec l'automatisation de la production de nouvelles formes graphiques.

Dans cette perspective, l'image numérique a produit une première révolution avec l'émergence des images générées par ordinateur à partir de la modélisation 3D. Auparavant, les images produites de manière mécanique relevaient toujours de l'empreinte, telle que la définit Georges Didi-Huberman²⁷ : inscrite dans les arts mécaniques, celle-ci procède de pratiques visant à automatiser, et donc à réduire la part d'interprétation humaine dans la production d'images. Le moulage, l'empreinte, le tracé des ombres sont, comme la photographie, des outils de génération de formes conduisant à une inflation du nombre d'images dans le monde, de traces externes reposant sur une forme de transfert ou de transposition à partir d'une forme. Au contraire, les arts libéraux mettent au centre de leur pratique l'humain, seul capable de produire des images nouvelles, d'inventer ou de recomposer des formes.

Les images générées par ordinateur, en particulier celles des mondes virtuels du type *Second Life* ou des films d'animation Pixar, ont provoqué de vives critiques (négatives comme positives) lorsqu'elles ont commencé à être produites massivement dans les années 1990, précisément parce qu'elles ne relevaient plus du paradigme de l'empreinte : ce qui choquait était l'absence de tout référent, la production purement mathématique d'espaces et de représentations imitant le réel²⁸. On peut pourtant faire remonter les principes de production de ces images aux procédés de mathématisation du regard : le calcul de la perspective a permis aux peintres de produire des images en fonction d'une théorie de la vision (en partie arbitraire), et d'en écarter d'autres : on a mis en place un espace visuel autorisé, existant dans un cadre mathématique précis²⁹.

Mais depuis peu (le début des années 2020), les progrès des outils de traitement automatique des images par apprentissage profond ont permis l'apparition de nouveaux processus de génération d'images reposant sur des bases très différentes. L'essor actuel des outils de production d'images à partir de

27. Georges Didi-Huberman, *La ressemblance par contact : archéologie, anachronisme et modernité de l'empreinte*, Paris, Éditions de Minuit, 2008.

28. Voir W.J.T. Mitchell, *The Reconfigured Eye: Visual Truth in the Post-Photographic Era*, Cambridge/Londres, MIT Press, 1992.

29. Michael Baxandall, *L'œil du Quattrocento*, Paris, Gallimard, 1985.

texte (*text-to-image*, comme Midjourney, Dall-E 2, Stable Diffusion, Imagen) a conduit des milliers de personnes à explorer les potentialités de production d'images synthétisées dans des modèles pré-entraînés, en partageant des conseils pour produire des images à partir de *prompts* (des descriptions textuelles) originaux qui prennent en compte les particularités syntaxiques et les métadonnées associées aux images pour générer des objets selon des styles particuliers³⁰ (la Cène dans le style de Basquiat par exemple).

Ces modèles, qu'on peut considérer comme des mémoires artificielles de la culture visuelle humaine, ont été produits par la vectorisation de millions d'images et leur appariement avec des vecteurs de mots liés à leurs légendes ou leurs métadonnées. Autrement dit, dans un modèle comme Stable Diffusion, une masse d'images et de métadonnées, représentant une part conséquente des images disponibles dans des bases de données plus ou moins représentatives d'une culture visuelle (humaine) communément partagée, a été réduite à des coordonnées dans un espace vectoriel complexe. La distribution des données ainsi calculée constitue un « modèle de diffusion latente » qui traduit les données visuelles et textuelles encodées en un espace de représentation potentiel³¹. À partir de ce modèle, on peut dès lors partir d'un texte (un *prompt*) pour contrôler la génération d'une image que le modèle « hallucine³² ».

On peut concevoir les modèles ainsi entraînés comme des formes de cartographies : ils simulent une partie de la culture visuelle de nos sociétés en synthétisant des pratiques et des objets à une très grande échelle, et permettent de l'explorer comme un espace virtuel³³. Pour penser ces nouvelles formes de création, François Chollet, à l'origine de l'un des outils les plus utilisés de *machine learning*, Keras, et l'un des principaux théoriciens du champ de l'intelligence artificielle, propose de changer de métaphore : plutôt que de considérer ces productions comme des créations, il faut les voir comme une nouvelle forme de photographie : « La génération d'images est une forme de

30. Jonas Oppenlaender, « A Taxonomy of Prompt Modifiers for Text-To-Image Generation », *arXiv*, 2022. DOI : 10.48550/arxiv.2204.13988 ; Vivian Liu et Lydia B. Chilton, « Design Guidelines for Prompt Engineering Text-to-Image Generative Models », dans *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 29 avril 2022, p. 1-23. DOI : 10.1145/3491102.3501825.

31. Ted Underwood, « Mapping the Latent Spaces of Culture », *Startwords*, n° 3, 2022. DOI : 10.5281/zenodo.6567481.

32. Robin Rombach *et al.*, « High-Resolution Image Synthesis with Latent Diffusion Models », *arXiv*, 13 avril 2022. DOI : 10.48550/arXiv.2112.10752 ; François Chollet, Luke Wood et Divam Gupta, « Keras Documentation: High-Performance Image Generation Using Stable Diffusion in KerasCV », *Keras*, 2022. En ligne : keras.io.

33. Estelle Blaschke, « From Microform to the Drawing Bot: The Photographic Image as Data », *Grey Room*, n° 75, 2019, p. 60-83. DOI : 10.1162/grey_a_00270.

photographie, dans un espace latent qui interpole entre des centaines de millions d'images. Lorsque vous prenez une photo, vous ne la "créez" pas, vous la prenez. Vous trouvez la scène que vous voulez, et vous la capturez comme vous le voulez. Il s'agit d'une forme de sélection³⁴. » Comme un photographe se promenant dans le monde, les utilisateurs de ces nouveaux outils explorent un espace mathématique, et prennent avec leurs *prompts* des instantanés de certains points d'intérêt dans cet univers virtuel. Le talent consisterait donc, comme pour la photographie, dans la dextérité à manier ces outils et à sélectionner les clichés les plus intéressants – exactement comme un photographe choisit parmi les milliers d'images qu'il a prises celles qu'il considère comme les plus marquantes, sans qu'on l'accuse de tricher, même s'il utilise le monde autour de lui plutôt que d'inventer ses images de toutes pièces.

D'une certaine manière, c'est ce que nous avons toujours fait, avec des moyens plus artisanaux. Toute création revient à élaborer une œuvre selon un ensemble de contraintes correspondant à un « point » de l'espace des possibles artistiques – l'originalité pouvant être définie comme une capacité à explorer des niches obscures de cet espace. Du côté de la production textuelle, l'idée maintes fois analysée de la « Bibliothèque de Babel » de Borgès, sorte de narrativisation du mythe de l'armée de singes pouvant produire tous les textes du monde en tapant de manière aléatoire sur des machines à écrire, exposait précisément ce rêve de voir incarner l'espace des possibles et de s'y déplacer pour découvrir, sans avoir à les produire, toutes les œuvres littéraires scriptibles. Alfred Jarry, dans *Gestes et Opinions du Docteur Faustroll, pataphysicien*, imaginait dès 1898 une « Machine à Peindre Clinamen » capable d'éjaculer des toiles de manière aléatoire, et donc de découvrir du nouveau, des formes visuelles échappant aux traditions humaines. Mais les espaces latents construits par les outils d'intelligence artificielle sont beaucoup plus restreints que les espaces infinis imaginés par Borgès ou Jarry – et c'est précisément ce qui fait leur puissance et définit leurs limites. Entraînés sur des millions d'images (dont certaines sont présentes des milliers de fois, comme les œuvres les plus reproduites du panthéon de la peinture occidentale), ces modèles ont construit des espaces gigantesques mais hiérarchisés et structurés, synthétisant des siècles de production et de visionnage pour y tracer des lignes de force selon les styles, les compositions, les choix de couleurs, les sujets de générations

34. François Chollet, dans son compte twitter, 2 septembre 2022 : « *Image generation is a form of photography. Photography in a latent space that interpolates between hundreds of millions of images. When you take a photo, you don't "create" the picture, you take it. You find the scene you want, and you capture it the way you want. It's curation.* » Nous traduisons.

d'artistes et de producteurs d'images – dans le cadre des limites que les créateurs de ces modèles ont imposées (filtres sur les contenus pornographiques, gores ou racistes, surreprésentation de l'imagerie occidentale, correction de certains biais de représentativité des minorités). Les résultats de l'exploration de ces espaces seront toujours intéressants, parce que, même « ratés », ils s'inscriront dans le champ de nos représentations (des représentations qui ont largement fait preuve de leur efficacité sur nos appareils cognitifs) ; mais ils ne pourront explorer que des espaces déjà connus, qui auront servi à leur entraînement. Des critiques se sont déjà élevées pour écarter ces nouvelles images, condamnées selon leurs détracteurs à ne faire que répéter le connu et interdire l'inventivité. Élaborées à partir d'une vision traditionnelle de la création (comme puissance géniale de production *ex nihilo*), ces considérations sont en partie fondées, mais elles négligent la réalité des processus de création : les images que l'on voit désormais pulluler sur les réseaux sociaux, produites par des internautes, l'ont été au travers d'une collaboration entre l'homme et ces outils. Les choix de *prompts* (de plus en plus assistés eux-mêmes par des outils de génération de descriptions³⁵), la sélection opérée par les acteurs sur les images produites, leurs éventuelles retouches manuelles, puis leur validation collective (selon qu'elles sont reproduites, qu'elles circulent, qu'elles génèrent des partages ou des réactions) agissent comme des mécanismes de curation de ces nouvelles images. Les modèles graphiques d'IA s'insèrent dans une culture visuelle préexistante qu'ils contribuent à transformer et qui agit en retour sur les images produites.

Ces nouvelles formes de génération d'images soulèvent des débats légaux et moraux inédits, et on voit les collectifs d'artistes et les industries culturelles obligés de réagir, et de nouvelles pratiques collectives se mettre en place, qui deviendront sans doute assez rapidement des standards puis des lois. Une grande quantité d'images utilisées pour entraîner ces modèles sont des images sous droits, qui ont été utilisées sans l'autorisation de leurs auteurs. Ces images disparaissent entièrement dans le processus d'entraînement, et ne peuvent être retrouvées (sauf cas extrême, comme des images très célèbres) dans le modèle lui-même ; mais il suffit d'utiliser dans les *prompts* de génération le nom d'un auteur pour produire des images imitant son style, pour peu qu'un nombre suffisant d'images ait été utilisé durant l'entraînement. Ceci permet de créer des œuvres inédites de/selon Gauguin, Miro, Giotto sur les sujets les plus divers ; ce numéro est ainsi illustré par des variations autour de la notion d'« œil numérique » par de grands noms de l'histoire de l'art. Mais le procédé peut

35. Voir par exemple : <https://promptomania.com/>.

être utilisé pour des artistes encore vivants : les communautés de graphistes, craignant pour leurs revenus, se sont rapidement prononcées collectivement pour bannir toute utilisation commerciale des productions de ce type. On voit également resurgir des débats sur les questions d'auctorialité, de maîtrise technique, d'originalité suscités en leur temps par les *ready-made* ou d'autres pratiques comme le *dripping* : lorsqu'un artiste a gagné un concours d'illustration numérique avec une peinture générée par une IA, mérite-t-il sa première place³⁶ ? Le cas le plus problématique reste cependant celui d'un internaute ayant « fini » la peinture numérique d'un artiste avec ces outils, alors même que l'artiste était en train de diffuser en direct son travail : il a suffi au faussaire de faire une capture d'écran de la peinture en cours et de remplir les espaces encore vides grâce aux outils de génération automatique de Dall-E pour en proposer une version terminée avant l'artiste. Le plagiaire a été jusqu'à poster son image sur le fil de discussion de l'artiste en l'accusant de l'avoir copié, et en exigeant d'être mentionné comme inspiration pour son œuvre, ce qui n'a pas manqué de susciter l'indignation des followers de l'artiste et plus largement du monde de la création numérique³⁷. On voit ici que les protagonistes endossent des comportements habituels aux artistes, et jusqu'à un certain point leurs (mauvaises) pratiques – c'est ici le mode créatif et surtout réactif qui modifie la donne.

Qu'est-ce que cela nous apprend sur la nature même de ces images ? Tout simplement qu'il s'agit moins d'objets que de potentialités au croisement de contraintes matérielles (les organes visuels, les outils techniques qui permettent de les produire et dont on peut saisir les effets en tapant le nom dans les *prompts*), plastiques (la surface plane et les jeux de couleurs), symboliques (l'organisation des images liée à des hiérarchies et structures sémantiques), dont les réalisations actuelles ne sont qu'une partie infime des productions possibles. L'analyse de ces dernières comme un chemin (choisi) dans ce vaste espace potentiel, plutôt que comme une histoire linéaire ou fondée sur l'idée de progrès technique ou esthétique, permet de repenser l'histoire culturelle³⁸.

36. Kevin Roose, « An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy », *The New York Times*, 2 septembre 2022.

37. Sisi Jiang, « Thief Steals Genshin Impact Fan Art Using AI, Demands Credit From Creator », *Kotaku*, 13 octobre 2022.

38. Baptiste Caramiaux et Sarah Fdili Alaoui, « "Explorers of Unknown Planets": Practices and Politics of Artificial Intelligence in Visual Arts », dans Association for Computing Machinery, *Proceedings of the ACM on Human-Computer Interaction*, 2022. DOI : 10.1145/1122445.1122456 ; X. Y. Han *et al.*, « Chapter 1: Artificial Intelligence and Discovering the Digitized Photoarchive », dans Lise Jaillant (dir.), *Archives, Access and Artificial Intelligence: Working with Born-Digital and Digitized Archival Collections*, Bielefeld, Bielefeld University Press, 2022, p. 29-60. DOI : 10.1515/9783839455845-002.

L'agentivité de l'œil numérique : au-delà de l'anthropologie des images

L'œil numérique peut ainsi désigner l'ensemble des médiations nouvelles qui créent une culture visuelle devenue hybride, puisque la cognition est partagée et répartie (distribuée) dans des systèmes numériques entre agents cognitifs humains et non-humains. Cette forme de regard est comme toute vision un phénomène historique. Fondé sur la virtualisation des images permettant une médiatisation algorithmique de leur production et de leur diffusion, cet œil numérique est un nouveau principe d'organisation du visible, qui transforme notre culture visuelle et, plus profondément, modifie notre compréhension de ce qu'est une image. On peut en appeler à une iconologie, comprise non plus comme une série d'images construites par la critique, mais comme l'existence en série des images qui réduit l'importance de chaque image singulière et conduit à examiner les hyperobjets que sont les images en réseau.

L'œil numérique pourrait en définitive désigner cette nouvelle agentivité qui a reconfiguré le monde des images, autrement dit nos écosystèmes visuels en articulant ces divers mécanismes : une forme de cognition algorithmique décentralisée et ubiquitaire qui se surajoute à nos processus sociaux, institutionnels, techniques de sélection et de production d'images, et qui agit non seulement au niveau des objets mais aussi au niveau des contenus symboliques et de leur visibilité. L'œil numérique représente la délégation d'une partie des outils de traitement de l'imagerie à des formes d'automatisation cognitives. Il ne s'agit plus de prothèses, mais de « collaborateurs » dans la configuration d'une culture visuelle rénovée. On peut d'ailleurs se demander si une partie de l'art numérique, qui ne repose que partiellement sur la computation mais engage également des objets physiques, ne vise pas à restaurer la primauté d'une création – essentiellement (au sens fort) – humaine. L'« anthropologie des images », développée par Hans Belting en partant du principe que « ce sont les hommes qui ont fabriqué et qui continuent encore de fabriquer des images³⁹ », est ébranlée dans ses fondements mêmes. Nous ne sommes plus les seuls à voir et à générer les images.

39. Hans Belting, *Pour une anthropologie des images, op. cit.*, p. 8.