

Des visages invisibles. Vers une histoire des images de la reconnaissance faciale

« Rien n'est moins personnel que le visage », c'est ce qu'on peut lire dans un passage des *Dialogues* entre Gilles Deleuze et Claire Parnet parus en 1977.¹ Une telle phrase résonne de manière familière à l'oreille des lecteurs et lectrices de *Mille plateaux*. En effet, les *Dialogues* anticipent certains aspects de ce que Deleuze et Félix Guattari développeront dans l'un des chapitres les plus fréquemment cités de leur *opus magnum* qui paraît trois ans plus tard.² Rien n'est moins personnel que le visage, ce qui signifie déjà ici : le visage n'est pas une partie du corps humain naturellement donnée, mais un vecteur de signes et d'informations socialement produit, projeté pour ainsi dire par une « machine abstraite »³ sur la face avant de la tête. Ce visage désincarné apparaît dans un premier temps comme un motif visuel bidimensionnel constitué de taches plus ou moins claires et foncées, qui se voit ensuite pourvu de caractéristiques identifiables à l'issue d'une suite d'opérations de décisions binaires multi-étapes avant d'être classé au moyen d'une grille de différences et situé dans un espace de « visages élémentaires » stéréotypés.⁴ Partout où apparaît un visage, il n'est pas seulement objet, mais aussi produit de ces processus d'analyse combinatoire et d'une reconnaissance des formes, qui opèrent de manière quasi automatisée et qui visent à attribuer aux individus une place dans le tissu social, et à les rendre identifiables, pour ne pas dire maîtrisables. À l'inverse, dans *Mille Plateaux*, Deleuze et Guattari se concentrent sur la fonction déterritorialisante du fait de « défaire le visage » et de fuir dans la non-reconnaissabilité.⁵

Lorsqu'on lit ces descriptions d'une production « mécanique » des visages, près de quarante ans après leur rédaction, elles semblent de prime abord d'une incroyable actualité. *Rien n'est moins personnel que le visage* vaut aujourd'hui plus que jamais : à l'époque des caméras omniprésentes, de flux d'images circulant de manière mondialisée et de la reconnaissance faciale automatisée, les visages apparaissent de plus en plus comme des profils d'informations désincarnés, voués à être saisis et traités par des machines. Depuis le *smartphone* déverrouillé au moyen de l'identification faciale, jusqu'aux gigantesques banques de données de visages constituées par nombre de services de renseignements, d'autorités de sûreté et d'entreprises privées : ce sont de plus en plus souvent des algorithmes fouillant les flux de données à la recherche de schémas récurrents qui déterminent ce qu'est un visage et à qui il appartient.

Ce n'est pas un hasard si les technologies actuelles de la reconnaissance faciale automatisée ont parfois été interprétées comme une « manifestation » concrète de la

« machine abstraite » de Deleuze et Guattari.⁶ Mais une telle interprétation, pour séduisante qu'elle soit, omet de prendre en compte deux aspects. Elle néglige d'une part l'historicité de la théorie de la visagité de Deleuze et Guattari, car même si leur projet théorique renvoie à un horizon dépassant de telles contingences temporelles, son apport spécifique serait incompréhensible sans l'horizon politique, technique et épistémique des années 1970. Ce sont précisément ces années qui ont vu à la fois une radicalisation des combats politiques concernant l'identification et l'émergence d'une conjoncture favorable à la recherche sur la reconnaissance faciale automatisée dont on peut reconnaître des échos dans les descriptions de la « machine abstraite ».

D'autre part, l'analogie sans solution de continuité entre « machine abstraite » et reconnaissance faciale algorithmique conduit à occulter le fait que cette dernière s'est profondément transformée, depuis ses débuts, en termes de fonctionnement tant technique que social : si l'on parlait encore du principe dans les années 1970 que la perception humaine des visages et le traitement automatisé des données pouvaient être comparés dans leur principe, voire même mis en équivalence sur le plan structurel, les algorithmes actuels de reconnaissance faciale se fondent sur des processus de traitement statistique automatisé qui n'ont plus rien en commun avec la façon dont des sujets humains se reconnaissent entre eux. L'emploi d'éléments théoriques issus du post structuralisme, simplement transposés au temps présent, nous rend non seulement aveugles à l'historicité de la théorie, mais méconnaît également la spécificité historique de la situation actuelle.

L'une des thèses de cet essai est la suivante : on ne peut clairement saisir ce que signifie aujourd'hui la reconnaissance faciale automatisée que lorsqu'on se représente ce qu'elle a signifié au cours de son histoire de plus de cinquante années. Cette histoire, dont nous retraçons les étapes décisives dans la partie principale de cet essai, est aussi une histoire de l'image, retraçant l'évolution des représentations quant au potentiel des images à révéler les visages, quant à la façon de les exploiter d'un point de vue technique et au type d'informations générées à partir d'elles et quant à la manière dont s'entremêlent la perception subjective et la pratique culturelle des images avec le traitement automatique des données. Et le projet théorique de Deleuze et Guattari, bien que ne faisant pas partie de cette histoire des images, se laisse lui aussi tout à fait remettre en perspective face à cet horizon historique.

1977 — DEVANT UN MUR BLANC

Remontons donc tout d'abord à l'année 1977 et aux *Dialogues* entre Deleuze et Parnet. Les événements de cette année mouvementée ont en effet imprimé leur marque de manière sous-jacente à la théorie de la visagité.⁷ On trouve notamment au cœur de ces *Dialogues* la célèbre formule du « système mur blanc – trou noir » que Deleuze y introduit comme produit de la convergence d'intérêts théoriques appartenant à l'origine à des champs différents. Tandis que Guattari est fasciné par le « trou noir » de l'astronomie, comme attracteur ultime, « ce qui vous capte et ne vous laisse pas sortir » ; Deleuze apporte quant à lui le concept de « mur blanc » qui, compris comme une toile vierge et comme surface de projection, lui permet de tisser des liens avec

des problématiques relatives à la peinture et au cinéma.⁸ Combinés entre eux, ces deux concepts constituent un schéma binaire de la visagité élémentaire :

« Car des trous noirs sur un mur blanc, c'est précisément un visage, large visage aux joues blanches et percé d'yeux noirs, ça ne ressemble pas encore à un visage, c'est plutôt l'agencement ou la machine abstraite qui va produire du visage. »⁹

On pourrait ici, en s'appuyant sur Sigrid Weidel, penser à l'une des célèbres photographies de la série des *Graffiti* de Brassäi (Ill. 1), qui met en image *La Naissance du visage* à partir de la rencontre apparemment fortuite entre des failles, des fissures et des taches sur les murs de Paris.¹⁰ Cependant, dans les explications de Deleuze au sujet de la fonction concrète du « système mur blanc – trou noir », d'autres associations plus actuelles et bien plus politiques semblent jouer un rôle :

« C'est cela, être identifié, fiché, reconnu : un ordinateur central fonctionnant comme trou noir et balayant un mur blanc sans contours. »¹¹

Dans l'image des murs blancs telle qu'elle est évoquée ici, on peut reconnaître sans peine les murs contre lesquels la police adosse ceux dont elle veut enregistrer les données personnelles, qu'elle veut photographier ou identifier lors d'une confrontation. Et le trou noir reflète autant l'œil sans âme de l'objectif photographique que « littéralement »¹² le « trou », donc la geôle dans laquelle les prévenus finissent par atterrir.

Qu'est-ce que tout cela a à voir avec l'année 1977 ? Pour Deleuze comme pour Guattari, l'année 1977 est une année au cours de laquelle les recherches et les arrestations, les conflits avec la police et la justice sont au cœur des débats politiques – à côté des arrestations de Franco « Bifo » Berardi et de Toni Negri en Italie, ont avant tout lieu les événements liés à l'« automne allemand ». Devant la menace imminente d'une arrestation, Klaus Croissant, l'avocat qui assure la défense d'Andreas Baader lors du procès Stammheim, s'enfuit à l'été 1977 en France où il demande l'asile politique. Les autorités allemandes l'accusent de soutien à une organisation terroriste et en septembre, au point culminant de l'« automne allemand », Croissant est arrêté à Paris. Quasiment toute l'intelligentsia française de gauche proteste contre son extradition (qui aura tout de même lieu en novembre). Michel Foucault, ainsi que Deleuze et Guattari, prennent part en première ligne à ces protestations. Mais tandis que Foucault veut s'engager dans la défense d'un avocat contre des répressions politiques, tout en rejetant les actions terroristes, Deleuze et Guattari, voyant dans cette affaire les signes annonciateurs d'une « dictature policière » outre-Rhin, ne font pas mystère de leur sympathie avouée pour la lutte armée.¹³ Quand Deleuze évoque donc le fait d'afficher et d'identifier, il pourrait alors fort bien s'agir d'une allusion aux avis de recherche des membres de la RAF qui couvrent les murs de toute la République fédérale.¹⁴ Car durant les « années de plomb », il n'y a pas qu'en Italie et en RFA¹⁵ que la peur des mesures de recherche omniprésentes et des arrestations arbitraires fait partie de la réalité quotidienne des intellectuels de gauche. C'est ainsi que Foucault est notoirement plaqué dos au mur et fouillé par une bonne douzaine de policiers lourdement armés alors qu'il quitte son hôtel à Berlin-Ouest – et ce, uniquement parce que quelqu'un a confondu le visage de son accompagnatrice, Heide Paris, éditrice chez Merve, avec celui de Inge Viett, pour laquelle un avis de recherche avait été publié.¹⁶

En 1977, ces visages sur des murs blancs, affichés et identifiés par des appareils anonymes ne sont pas qu'une métaphore anhistorique, mais font bien partie d'une réalité politique concrète dans laquelle la circulation d'images (via les avis de recherches) décide du destin d'êtres humains. L'« ordinateur central » est même déjà en service au sens très « littéral » du terme, puisqu'à Wiesbaden en particulier, Horst Herold, le chef de la police criminelle fédérale que la presse n'appelle plus que du nom de « commissaire ordinateur » a mis au point le système de traitement de données à des fins policières le plus moderne d'Europe. Certes l'appareil de recherches hautement équipé d'outils numériques qu'il a développé ne dispose pas encore d'une reconnaissance faciale automatisée, mais de telles technologies ne sont à l'époque déjà plus de la pure science-fiction. Aux États-Unis et au Japon, nous verrons que cela fait déjà plus d'une décennie que la recherche explore le sujet et la police criminelle de Wiesbaden espère elle aussi que le développement de « systèmes électroniques d'identification des personnes » et de banques de données de visages aboutira à la fin des années 1970 à l'optimisation de l'interpellation de suspects en cavale.¹⁷ Même si Deleuze et Guattari n'avaient peut-être pas directement connaissance de ces recherches appliquées, les fantasmes liés à une automatisation de la vision sur laquelle celles-ci sont fondées, auraient tout à fait pu leur être familiers, notamment par la littérature secondaire de l'époque en psychologie. Car le fait de ramener la reconnaissance automatique de formes à la perception humaine des visages est un projet aussi bien technique que théorique, pour lequel non seulement les criminalistes et les informaticien·ne·s, mais aussi les chercheur·e·s en sciences cognitives et les psychologues montrent un vif intérêt dès les années soixante.¹⁸

1970 — DES MATRICES LISIBLES PAR DES MACHINES

L'histoire commence tout juste à lever le voile sur les débuts de ce projet qui restent pour une grande part encore obscurs.¹⁹ Les tentatives d'apprendre aux ordinateurs à « voir » remontent en réalité à la première phase de recherche intensive sur l'Intelligence Artificielle dans les années cinquante. Mais les succès pratiques restent rares et se limitent à des domaines spécifiques comme la reconnaissance en écriture. Celle-ci est néanmoins déjà suffisamment avancée dès 1959, tout du moins en ce qui concerne les écritures « lisibles par des machines » spécialement développées pour l'occasion, afin que la *Bank of America* puisse commencer à traiter automatiquement les chèques.²⁰ Les visages humains en revanche ne s'avèrent être « lisibles par des machines » qu'à certaines conditions. C'est ce dont le mathématicien américain Woody Bledsoe, qui peut être considéré comme pionnier de la reconnaissance faciale automatisée, doit faire l'expérience.²¹ Après avoir développé dès la fin des années cinquante des méthodes tout à fait innovantes de reconnaissance de formes et plus spécifiquement d'écritures, il se met à son compte et fonde son entreprise *Panoramic Research Inc.* en 1960 à Palo Alto. Parmi ses clients, il compte, outre le Ministère de la Défense américain, la CIA, par laquelle il est missionné en 1963 pour entamer des recherches sur la reconnaissance faciale automatisée. Le projet est bien plus ambitieux que la reconnaissance scripturale, malgré les succès partiels de celle-ci, car les visages se distinguent

à plus d'un égard des lettres de l'alphabet latin. Ils n'existent pas seulement en 26 versions, mais en un nombre quasiment infini de variantes et toutes ces variantes reposent sur des combinaisons des mêmes éléments récurrents : yeux, nez et bouche. En outre, les visages ne sont pas des formes graphiques discrètes, stables et planes, mais des structures continuellement vivantes, mouvantes et tri-dimensionnelles, dont la forme visible se modifie en permanence en fonction de l'âge, de l'expression et de conditions variables d'éclairage. Bledsoe travaille pendant plusieurs années à l'élaboration d'un procédé qui doit venir à bout de ces défis, mais le mieux qu'il puisse proposer à ses commanditaires est un système semi-automatisé : des opérateurs humains marquent les points caractéristiques prédéfinis comme les coins des yeux et de la bouche, la pointe du nez et du menton sur des images standardisées de recherche policière avec l'aide d'une tablette graphique RAND. Les coordonnées ainsi saisies servent ensuite à l'identification par l'ordinateur qui les compare avec des ensembles de données déjà saisies en amont. Il s'agit par conséquent d'un « système homme-machine » avec une répartition très claire des tâches : seule la gestion de masses de données invisibles est confiée à l'ordinateur, tandis que la reconnaissance des formes dans le champ du visible se voit déléguée à des paires d'yeux humains. Autrement dit : l'ordinateur est dans ce dispositif expérimental primitif (encore) frappé de cécité.

Le travail de Bledsoe reste encore confidentiel plusieurs décennies durant, ce qui sert tout à fait les intérêts de son commanditaire dans l'ordre du renseignement, et ce n'est que très récemment qu'il est devenu accessible au grand public.²² La *Nippon Electric Company* (NEC), qui travaille également dans les années soixante sur les techniques de reconnaissance faciale automatisée, a quant à elle, un bien plus grand intérêt à ce que son travail soit connu du public. L'un des résultats spectaculaires au premier stade de sa recherche est présenté lors de l'exposition universelle de 1970 dans la ville japonaise d'Osaka, où l'on propose une attraction très spéciale au public sous le titre de *Computer Physiognomy* (Ill. 2). Tous les visiteurs et visiteuses de l'Expo acceptant de mettre leur visage à disposition du traitement électronique de données ont le privilège de remporter chez eux comme souvenir non seulement un portrait d'eux-mêmes pixellisé et imprimé, mais ils font également l'expérience d'une forme très spéciale d'interprétation automatisée des caractères. Pour ce faire, ils sont tout d'abord installés sur un tabouret rehaussé qui semble plutôt inconfortable et se voient littéralement fixés et isolés devant un « mur blanc ». Leur visage est ensuite enregistré par une caméra vidéo puis numérisé, avant d'être finalement apparié à l'un des sept visages de célébrités proposés à la comparaison sur la base de ressemblances mesurables. Chacune de ces célébrités, de Winston Churchill à John F. Kennedy et Marilyn Monroe, est censée représenter un certain « type », mais les visiteurs et visiteuses ne sont en rien informés des critères selon lesquels est établie cette affiliation. Bien plus, le résultat de cette « physiognomie par ordinateur » est présenté au public comme s'il était rendu par un oracle. Ce qui se joue à l'intérieur de l'ordinateur reste dans un premier temps également caché aux informaticien.ne.s prenant part à l'expérience, qui ne doivent apprendre que lors de l'analyse de données réalisée *a posteriori* que l'ordinateur travaillait avec un taux d'erreur très élevé et que les résultats n'étaient la

plupart du temps que le fruit du hasard.²³ La reconnaissance faciale, et cela vaut jusqu'à aujourd'hui, se présente souvent comme une *black box* et il n'est pas rare que les critères sur la base desquels elle opère restent structurellement cachés à ceux qui les mettent en œuvre, tout comme à ceux qui sont enregistrés par elle.

Malgré son immaturité sur le plan technique, le divertissement d'Osaka s'avère d'une importance capitale. Car sur la base des expériences acquises, et avant tout sur celle de la collecte des données, le jeune informaticien Takeo Kanade développe l'un des premiers procédés qui soit à peu près fonctionnel pour la reconnaissance faciale, indépendamment de l'aide humaine.²⁴ De même que le système homme-machine de Bledsoe en son temps, Kanade mise sur le relevé de points caractéristiques du visage et la mesure de la distance entre ceux-ci. La condition préalable à cela est un processus de réduction visuelle radicale et d'abstraction graphique (Ill. 3). Le visage photographié de face, isolé de manière artificielle devant un mur blanc et enregistré par vidéo, comme à Osaka, est dans un premier temps transformé en une matrice de pixels en noir et blanc, dont on n'isole et conserve que dans un second temps que les contours. Dans cette configuration linéaire discrète, l'ordinateur localise et quantifie ensuite des grandeurs mesurables prédéfinies, par exemple l'écart entre les yeux. Comme base de cette saisie automatisée d'éléments caractéristiques, on se sert d'un schéma général de la visagité qui renseigne le programme sur l'endroit probable de la matrice de pixels où l'on peut trouver un œil, une bouche ou un nez. Dans une certaine mesure, l'ordinateur « sait » donc en amont à quoi ressemble un visage en général et saisit en partant de cette base les combinaisons particulières et respectives entre œil, bouche et nez. Le procédé est toutefois susceptible de commettre des erreurs – des contours supplémentaires dans le visage, comme ceux que créent les rides, les barbes ou les lunettes, rendent la reconnaissance de caractéristiques quasiment impossible. Aussi Kanade se concentre-t-il sur des visages jeunes, dépourvus de barbe et de lunettes et qui plus est, uniquement sur des visages masculins de type japonais.²⁵ Le schéma général de visagité sur lequel repose son système se définit comme universel, mais il n'est en aucun cas inclusif. La reconnaissance faciale, et ceci vaut jusqu'à aujourd'hui, opère constamment avec de tels standards explicites comme implicites : tous les visages ne sont pas également lisibles par les machines.²⁶

INTERLUDE AU TOURNANT DES ANNÉES 1980 — HUMAIN ET HOMME-MACHINE

Ces premiers pas dans la reconnaissance faciale automatisée prennent place dans la longue tradition de la mesure et du relevé de données anthropométriques, dont on peut suivre la trace jusqu'au XVIII^e siècle. Le *Computer Physiognomy* porte bien son nom, sans doute plus que ce que pouvaient l'imaginer ses développeur·se·s japonais·e·s. Le pasteur suisse Johann Caspar Lavater, qui fit de l'interprétation physiognomonique des caractères une mode dans toute l'Europe dans les années 1770, inventa en son temps avec la machine à silhouettes un dispositif spatio-médial conçu pour isoler les traits du visage humain, les stabiliser et les saisir avec une exactitude toute mécanique (Ill. 4). Et la technique d'ombre projetée élaborée par Lavater servait déjà à extraire les contours du visage (ou plus exactement du profil), pour les décomposer ensuite en

distances mesurables et en formes équivalentes, afin de pouvoir les lire à la manière d'un alphabet.²⁷ Le *signalment* anthropométrique du criminaliste parisien Alphonse Bertillon élaboré dans les années 1880, dont Woody Bledsoe se servit encore explicitement comme d'un modèle, peut lui aussi être placé dans la droite ligne de cette tradition.²⁸ Indépendamment du fait de savoir à quel point ces procédés sont liés à l'histoire du colonialisme et du racisme européen, la façon dont ils s'emparent des corps et des visages en les mesurant et en les morcelant, paraissait déjà à leurs protagonistes être à l'exact opposé de la façon dont fonctionne la perception humaine des visages au quotidien. Car nous reconnaissons habituellement les visages non par des caractéristiques isolables et des distances mesurables, que nous comparons en nous conformant à des règles explicites, mais plutôt dans leur globalité formelle, que nous enregistrons de manière intuitive.²⁹ Et tandis que les humains reconnaissent encore un visage familier après de nombreuses années et même dans les circonstances les plus défavorables, la moindre barbe ou une simple paire de lunettes vient mettre la mesure électronique en échec. Le fait de comprendre cette aptitude humaine et de la simuler par ordinateur n'offrirait-il pas le chemin le plus prometteur menant à l'automatisation de la reconnaissance faciale ?

En effet, parallèlement aux premières expériences en matière de reconnaissance faciale automatisée, la recherche fondamentale en psychologie se tourne dans les années 1970 et 1980 elle aussi vers la perception des visages. Ce faisant, elle garde en ligne de mire la transposition de l'homme à la machine, au moins comme horizon lointain. Car depuis le « tournant cognitif », on est tout à fait convaincu que le mode de fonctionnement de la perception humaine des visages se laisse *in fine* décrire comme une sorte de programme informatique, qui encode un *input* sensoriel selon des règles explicites avant de le stocker, de le récupérer et de le transformer.³⁰ La force de séduction de cette analogie entre l'homme et la machine réside dans la promesse d'un transfert théorique réciproque : l'exploitation électronique des données fournit un modèle pour des processus cognitifs en principe invisibles et l'on espère que les connaissances tirées de la psychologie expérimentale pourraient contribuer désormais à l'amélioration d'algorithmes.³¹ La perception des visages est un traitement d'informations discrètes : voilà la formule aussi simple que pleine de présupposés à laquelle on peut résumer le consensus qui régit la recherche tant fondamentale qu'appliquée au tournant des années 1980. Et là où il s'agit de décrire des processus de traitement d'information, il semble d'une moindre importance de savoir s'ils se déroulent dans des cerveaux humains ou dans des processeurs électroniques – ils se laissent toujours modéliser sous forme de schéma abstrait, d'arbre décisionnel ou de logigramme complexe (Ill. 5).

C'est cet héritage de la cybernétique que Deleuze et Guattari semblent employer de manière critique, voire même par endroits détourner à des fins parodiques au travers du vocabulaire technique de leur chapitre sur la « visagété » : il y est non seulement question de manière récurrente de « zones de fréquence et de probabilité », d'« ordinateur central » et de « têtes chercheuses »,³² mais le concept de la « machine abstraite » en lui-même est issu de la cybernétique où il désigne une série d'étapes

de traitement discrètes qui peuvent ensuite techniquement être implémentées dans des formes concrètes variables, sous forme de programme ou de circuit de commutation par exemple.³³ Et en effet, des passages entiers du chapitre sur la « visagité » peuvent se lire comme la description sous forme de diagramme de la façon abstraite dont les programmes se déroulent, au cours desquels des visages sont décomposés en unités discrètes ensuite recombinaés pour former de nouvelles unités avant d'être sélectionnées successivement dans une suite d'opérations de décisions binaires.³⁴ Pourtant, là où les chercheurs et chercheuses désireux de décrypter la reconnaissance humaine des visages, croyaient être sur la piste d'un programme universel, Deleuze et Guattari ne laissent aucun doute quant au fait que leur « machine abstraite » est le produit d'une structure de pouvoir spécifique, à savoir de type « autoritaire » ou « despotique ». ³⁵ Le traitement de l'information, et cela reflète également la teneur concrète des discussions politiques à la fin des années 1970, apparaît ce faisant toujours comme une technique de domination. Et l'on comprend aussi de cette manière pourquoi la reconnaissance faciale signifie toujours aussi une production sociale des visages, car les visages en tant que vecteurs d'information lisible et identifiable ne sont jamais simplement donnés, et encore moins de manière naturelle, mais doivent être produits avec force moyens.

Le raccourci théorique entre reconnaissance humaine des visages et reconnaissance automatique de formes qui permet, autour de 1980, le transfert théorique entre les sciences cognitives et l'informatique, reste pour cette même raison une promesse illusoire. Car il est certes possible de décrire les processus de perception humaine comme des processus de traitement de l'information de manière spéculative, mais pas de les expliciter sous la forme de déroulements de programmes concrets : quant à savoir par le biais de quelles opérations concrètes on pourrait cheminer de l'« *input* sensoriel » à l'« information encodée », cela n'est aucunement explicité par les modèles abstraits de la perception humaine des visages, tout munis de leurs marqueurs et autres *black boxes* qu'ils soient (Ill. 5). L'analogie cybernétique entre l'homme et la machine s'avère être une métaphore attractive sur le plan théorique, mais conduit en pratique à une impasse : la période faste des années 1970 est ainsi suivie d'une décennie durant laquelle la recherche sur la reconnaissance faciale automatisée est dans les faits mise à l'arrêt, avec l'« hiver de l'IA » que constituent les années 1980.

1991 — ESPACES STATISTIQUES DU VISAGE

Un nouveau tournant est amorcé par une procédure qui ne se fonde plus sur l'analogie implicite entre la perception humaine et la reconnaissance automatique de formes, mais sur l'exploitation statistique des informations véhiculées par les images qui sont déjà à disposition sous forme de masses de données discrètes. Elle est présentée en 1991 par les chercheurs du MIT Matthew Turk et Alex Pentland, qui dès le début de l'essai « *Eigenfaces for Recognition* », qui sera bientôt le plus fréquemment cité dans ce champ de recherches, se démarquent très nettement de toutes les tentatives menées jusque-là en matière de reconnaissance faciale. Ces dernières seraient en effet parties de manière « intuitive » du fait que les visages se laissent le plus aisément

comparer et distinguer entre eux à l'aide de la géométrie formée par les yeux, le nez et la bouche.³⁶ Turk et Pentland en revanche prennent une autre voie : au lieu de chercher dans une image prise isolément des points caractéristiques définis au préalable, leur méthode s'appuie au départ sur plusieurs centaines de visages standardisés, à partir desquels le système apprend tout d'abord à reconnaître un visage. À partir de chacun des visages de ce qu'ils nomment un set d'entraînement, on établit une série de variations significatives par rapport à la moyenne de tous ces images faciales – ne réalisant cependant aucune mesure géométrique de distance entre des points isolés, mais recensant différents degrés de clarté répartis sur l'ensemble de la surface de l'image. Le résultat de cette décomposition analytique de l'image constitue un ensemble que l'on nomme des *eigenfaces* : des visages partiels schématiques, dont chacun représente certaines différences par rapport à la moyenne (Ill. 6). Chaque nouveau visage, auquel le système est confronté, est alors caractérisé comme « somme pondérée » de ces *eigenfaces*. Il est donc d'une certaine manière composé à partir d'éléments des visages d'entraînement et ne résulte que de la comparaison avec une archive numérique constituée de visages alternatifs. Chaque visage individuel apparaît ainsi comme l'actualisation concrète de visages possibles dans un espace certes virtuel mais multi dimensionnel et totalement prévisible.³⁷ La condition d'un tel procédé est un format uniforme des images : de même que pour certains procédés plus anciens, l'analyse des *eigenfaces* ne fonctionne, elle aussi, qu'à la condition que les visages soient présentés face caméra. Cela correspond tout à fait à leur contexte initial d'application. Cette méthode a été développée en collaboration avec un opérateur de taux d'audience télévisée qui voulait identifier la personne en train de regarder l'écran à l'aide d'une caméra fixée sur le téléviseur.³⁸

Certes Turk et Pentland soulignent bien les « similitudes qualitatives » supposées entre leur approche et la reconnaissance humaine des visages, sans doute en raison du contexte de publication de leur article dans le *Journal of Cognitive Neuroscience* qui relève des sciences cognitives, mais ils sont bien en peine de préciser en quoi elles consistent.³⁹ En effet, les abstractions statistiques avec lesquelles ils opèrent ne leur permettent aucune analogie de nature psychologique. Les *eigenfaces* représentent certes au sens mathématique des « caractéristiques » précises, mais justement pas de même nature que celles que des observateurs humains identifieraient comme des yeux, des nez ou des bouches. Il s'agit plutôt de proportions signifiantes de la répartition des valeurs qui échappent largement à toute description verbale. À la différence de l'écart entre les yeux par exemple, de telles caractéristiques ne pourraient pas non plus être lues à partir de photographies analogiques ou de corps vivants : il s'agit moins de caractéristiques isolées de visages individuels que de caractéristiques statistiques de masses d'images numériques. Une telle approche statistique fondée sur les images du visage humain n'est pensable que dans la mesure où Turk et Pentland, à la différence de Bledsoe et de Kanade, ne se fondent pas sur l'anatomie humaine, mais sur leur représentation en image, qui est toujours déjà disponible en tant que jeu de données numériques. Ainsi leur *pictorial turn*⁴⁰ de la reconnaissance faciale accomplit-il dans le même temps la rupture la plus radicale avec la perception humaine : la

reconnaissance faciale automatisée reconnaît désormais un visage non plus en vertu de configurations de caractéristiques visibles, mais de structures de données invisibles.

2014 — FLUX DES VISAGES INTERCONNECTÉS

L'analyse des *eigenfaces* passe pour être la première solution pratique au problème de la reconnaissance faciale. Elle déclenche non seulement un véritable boom de la recherche, mais éveille également l'intérêt d'institutions étatiques. Entre 1993 et 1996 le ministère de la Défense des USA invite chaque année dans le cadre du programme connu sous le nom de FERET (pour « Face Recognition Technology ») des chercheurs et chercheuses des universités d'élite américaines, pour tester les performances de leurs algorithmes de reconnaissance faciale en les soumettant à des tests standardisés. Les améliorations de la performance que l'on constate d'une année sur l'autre encouragent certains des participants à fonder leur propre start-up : au tournant de l'an 2000, plusieurs fournisseurs commerciaux de logiciels de reconnaissance faciale sont déjà en concurrence aux États-Unis pour remporter des marchés auprès des autorités de police ou des services de permis de conduire, ainsi que des exploitants de casino ou de stade. C'est enfin avec la « guerre contre le terrorisme » menée après le 11 septembre que cette jeune technologie – quoiqu'encore hautement faillible – quitte ensuite les laboratoires informatiques du complexe militaro-académique pour pénétrer dans la réalité mondialisée avec l'équipement « biométrique » accru des frontières et des passeports.⁴¹ Pour l'histoire de la reconnaissance faciale, le 11 septembre ne marque toutefois pas de césure décisive. Car les procédés qui sont à présent employés misent toujours sur des images largement standardisées, comme le montrent les conventionnelles « planches d'échantillons photographiques » de l'imprimerie fédérale, qui depuis 2005 visent à rendre largement publiques les consignes très strictes concernant le format de la photo d'identité du nouveau passeport électronique (Ill. 7), à savoir un fond neutre, un regard face caméra et surtout : aucun sourire !⁴²

Certes, la reconnaissance faciale au tournant de l'an 2000 opère avec la comparaison automatisée de grandes masses d'images, mais celles-ci sont structurées, formatées et *in fine* aussi limitées par la quantité de visages saisis numériquement. Cela change du tout au tout une décennie plus tard, avec la convergence de deux technologies jusque-là développées sans aucun lien entre elles, à savoir les « réseaux neuronaux artificiels » (RNA) et les réseaux « sociaux ». L'idée de simuler les processus neuronaux d'organismes vivants sur ordinateur remonte certes aux années 1940, mais ce n'est qu'autour de 2009 que les conditions semblent réunies pour espérer pouvoir atteindre des résultats convaincants dans le champ de la reconnaissance faciale. Indépendamment de leur appellation naturalisante, les réseaux neuronaux artificiels opèrent également avec le traitement statistique de grandes masses de données. Les critères de distinction et de classement qu'ils appliquent pour ce faire, ne leurs sont cependant pas prescrits par leur programmeur·rice·s. Au lieu de découper des images standardisées selon des règles fixes pour les réduire à leurs composantes statistiques, de tels algorithmes optimisent leurs procédures décisionnelles de manière autonome sur la base de processus itératifs d'optimisation : ils « apprennent » par des millions d'essais

et d'erreurs et, à cette fin, ils ont besoin de masses gigantesques de données pour s'entraîner.⁴³

C'est ici que les *social media* entrent en jeu, car pour s'entraîner à reconnaître un seul et même visage dans différentes circonstances, sous différents angles et dans des conditions d'éclairage variables, les algorithmes actuels sont nourris massivement avec des séries d'images qui doivent être attribuées à des personnes que l'on peut identifier de manière fiable – et que l'on appelle « *Faces in the Wild* », des visages qui n'ont pas été photographiés dans des conditions de laboratoire extrêmement standardisées, mais dans les conditions de la vie réelle.⁴⁴ Les masses gigantesques d'images de visages déjà « labellisés », c'est-à-dire reliés à un nom et à un profil ainsi que leur accessibilité sans difficulté font des plateformes telle que *Facebook* et *Instagram* des terrains d'entraînement et de test idéaux pour la reconnaissance faciale. Chaque fois que quelqu'un télécharge une photo de lui, reconnaît et « tague » une connaissance sur une photo ou vérifie une proposition faite par la reconnaissance faciale automatique, il ou elle contribue donc à l'entraînement des algorithmes.⁴⁵ Dès 2014, le département de recherche de *Facebook* peut se targuer d'avoir comblé « l'écart de performance » entre l'homme et la machine – leur algorithme *DeepFace* se montre à présent capable d'égaliser la faculté de reconnaissance humaine, tout du moins lors de tests standardisés (Ill. 8).⁴⁶ Toutefois, sans la contribution de milliards de paires d'yeux humains connectés aux réseaux mondiaux de circulation des images, les « réseaux neuronaux artificiels » resteraient aveugles, tout du moins jusqu'à aujourd'hui.

C'est pourquoi comprendre l'histoire de la reconnaissance faciale comme une histoire de l'image (*Bildgeschichte*) signifie prendre en considération le statut modifié des images dans les processus de traitement automatisé. Depuis son *pictorial turn* dans les années 1990, la reconnaissance faciale automatisée n'opère plus avec la fiction théorique du visage comme vecteur d'informations, dont les caractéristiques, visibles également pour des yeux humains, peuvent être exploitées dans des processus explicites de la reconnaissance de formes. Les algorithmes qui sont aujourd'hui employés dans les *black boxes* de notre présent numérique connecté, partent bien plus du principe selon lequel les données avec lesquelles ils opèrent sont toujours déjà disponibles en tant qu'ensemble de données d'images numériques – et, en cas de doute, que les informations qu'ils génèrent à partir d'elles ne sont pas celles que l'œil humain pourrait détecter face à un visage réel. La forme de production sociale des visages qui domine aujourd'hui reste mal comprise si on ne l'appréhende pas avant tout comme une production d'images partagée et connectée, autrement dit comme une production d'images dans laquelle les rôles de l'homme et de la machine apparaissent comme clairement définis et en aucun cas comme interchangeables.

Avec des réseaux sociaux de distribution d'images tels que *Facebook* et *Instagram*, ce sont des systèmes homme-machine globalisés qui vivent littéralement de ce que leurs utilisateurs ne se lassent pas de produire, de partager, d'évaluer et de commenter des images d'eux-mêmes et de tous ceux qui leur sont chers. Les images numériques de visages sont devenues une précieuse ressource de données pour des opérations d'apprentissage par la machine, le traitement statistique et la surveillance

automatisée. Cependant, là où elles ne sont que des bases de données invisibles et désincarnées pour les algorithmes qui, en arrière-plan de la plate-forme sociale, passent au crible les gigantesques flux d'images à la recherche de visages reconnaissables dans les coulisses des plateformes sociales, elles demeurent pour tous ceux à qui l'on n'offre rien d'autre à voir que la « face externe des données »,⁴⁷ leur face visible, leurs propres visages et ceux des êtres qui leur sont chers, ceux de personnes proches ou lointaines, qu'ils admirent ou envient, qu'ils gardent en souvenir ou préféreraient oublier. Chaque visage reconnaissable sur une image numérique sert ainsi à la fois de capteur d'attention humaine de nature affective et d'ancre opérationnelle permettant de relier des flux d'images de provenances diverses pour former des clusters de plus en plus vastes et des profils de plus en plus détaillés. Les images de visages sont ainsi devenues des interfaces qui garantissent la possibilité de relier les processus de traitement par la machine aux actes de la vision humaine. Visibles et invisibles à la fois, ces interfaces sont autant des schémas de données impersonnels que des surfaces de projection d'affects, d'émotions et de souvenirs tout à fait personnels.

- 1 Gilles Deleuze, Claire Parnet, *Dialogues*, Paris : Flammarion, 1996, p. 29.
- 2 Gilles Deleuze, Félix Guattari, « Année zéro. Visagéité » in : *idem*, *Mille plateaux : Capitalisme et schizophrénie*, Paris : Éditions de Minuit, 1980, p. 205-235.
- 3 *Ibid.*, p. 207.
- 4 *Ibid.*, p. 217-219.
- 5 *Ibid.*, p. 230.
- 6 Petra Löffler, « Effacement. Zur Auslöschung des Gesichts bei Gilles Deleuze », in : Mona Körte, Judith Elisabeth Weiss (éd.), *Gesichtsaufösungen (Interjekte 4)*, Berlin : ZfL, 2013, p. 87-96, ici : p. 89. C'est également l'hypothèse que fait l'une des études les plus approfondies jusqu'à présent des débuts de la reconnaissance faciale automatisée, voir Kelly Gates, *Our Biometric Future. Facial Recognition Technology and the Culture of Surveillance*, New York/Londres : New York University Press 2011, p. 23 et suiv.
- 7 C'est surtout pour Guattari, qui entretient des contacts étroits avec le mouvement autonomiste en Italie, que l'année 1977 est dans l'ensemble une année de combats politiques décisifs. Voir François Dosse, *Gilles Deleuze et Félix Guattari. Biographie croisée*, Paris : La Découverte 2007. Je remercie Ann-Cathrin Drews d'avoir attiré mon attention sur l'arrière-plan historique de l'époque.
- 8 Gilles Deleuze, Claire Parnet, *Dialogues*, *op. cit.*, p. 24.
- 9 *Ibid.*
- 10 Voir Sigrid Weigel, *Grammatologie der Bilder*, Berlin : Suhrkamp 2015, p. 150 et suiv.
- 11 Gilles Deleuze, Claire Parnet, *Dialogues*, *op. cit.*, p. 25
- 12 « Nous parlons littéralement. » *Ibid.*
- 13 Cette différence marquera *in fine* l'un des points de rupture de l'amitié entre Foucault et Deleuze, voir François Dosse, *Gilles Deleuze et Félix Guattari*, *op. cit.*, p. 727 et Didier Eribon, *Michel Foucault. 1926-1984*, Paris : Flammarion, 1989, p. 273-278. C'est également de ce dernier que provient la formulation « dictature policière » (*ibid.*, p. 276).
- 14 François Dosse rapporte que ce sont ces avis de recherche qui frappent, plus que toute autre chose, le peintre Gérard Fromanger, un ami commun de Deleuze et Guattari, lors de son voyage en 1977 en RFA. Voir François Dosse, *Gilles Deleuze et Félix Guattari*, *op. cit.*, p. 698.
- 15 *Ibid.*, p. 463-495.

- 16 Voir Michel Foucault, « Nous nous sentions comme une sale espèce » (1977), in *idem*, *Dits et Ecrits*, 1954-1988, tome III. 1976-1979, Paris : Gallimard, 1994, p. 415-418.
- 17 Voir Horst Herold, « Erwartungen von Polizei und Justiz in der Kriminaltechnik », in : Bundeskriminalamt Wiesbaden (éd.), *Der Sachbeweis im Strafverfahren. Arbeitstagung des Bundeskriminalamtes Wiesbaden vom 23. bis 26. Oktober 1978*, Wiesbaden : sans indication 1979, p. 75-83, ici : p. 79 et suiv.
- 18 Les développements suivants reposent en partie sur mon ouvrage *Operative Porträts. Eine Bildgeschichte der Identifizierbarkeit von Lavater bis Facebook*, Constance : Constance University Press, 2019.
- 19 Une histoire complète de la reconnaissance faciale automatisée manque encore dans la recherche. Gates en livre un premier aperçu dans *Our Biometric Future*, *op. cit.*, p. 25-61.
- 20 Voir Nils J. Nilsson, *The Quest for Artificial Intelligence. A History of Ideas and Achievements*, Cambridge et al. : Cambridge University Press 2010, p. 63 et suiv.
- 21 L'histoire de Bledsoe a récemment suscité un intérêt accru, voir Shaun Raviv, « The Secret History of Facial Recognition », in *Wired*, 22 janvier 2020, <https://www.wired.com/story/secret-history-facial-recognition/> [dernier accès 10/03/2020]. Voir en outre et pour les développements suivants : Anne Olivia Boyer, Robert S. Boyer, « A Biographical Sketch of W. W. Bledsoe », in : *idem* (éd.), *Automated Reasoning. Essays in Honor of Woody Bledsoe*, Dordrecht : Springer 1991, p. 1-30 ; Nilsson, *The Quest for Artificial Intelligence*, *op. cit.*, p. 127 et suiv. ; Gates, *Our Biometric Future*, *op. cit.*, p. 28-30.
- 22 Voir Woodrow Wilson Bledsoe, « A Proposal for a Study to Determine the Feasibility of A Simplified Face Recognition Machine », Palo Alto 1963 ; *idem*, « Facial Recognition Project Report », Palo Alto 1964, consultables tous deux en ligne : <https://archive.org/details/firstfacialrecognitionresearch> [dernier accès 10/03/2020].
- 23 Voir James L. Wayman, « The Scientific Development of Biometrics over the last 40 Years », in : Karl de Leeuw, Jan Bergstra (éd.), *The History of Information Security : A Comprehensive Handbook*, Amsterdam : Elsevier Science 2007, p. 263-274, ici p. 266 ; Gates, *Our Biometric Future*, *op. cit.*, p. 25 et suiv.
- 24 Takeo Kanade, *Computer Recognition of Human Faces*, Bâle/Stuttgart : Springer 1977. Le travail de Kanade a été soumis en 1973 à l'Université de Kyoto pour l'obtention d'un doctorat.
- 25 Voir *ibid.*, p. 10-13, 30-36, 66-69.
- 26 Ce problème est connu dans le débat actuel sous le nom de *algorithmic bias*, auquel Benjamin Ruha a livré récemment une introduction : Ruha Benjamin, *Race After Technology : Abolitionist Tools for the New Jim Code*, Cambridge : Polity 2019.
- 27 Voir à ce sujet Roland Meyer, « Netzwerke und Bilderströme. Johann Caspar Lavater als Medienunternehmer », in : Felix Lenz, Christine Schramm (éd.), *Von der Idee zum Medium. Resonanzfelder zwischen Aufklärung und Gegenwart*, Munich : Fink 2019, p. 319-337.
- 28 Voir Raviv, « The Secret History of Facial Recognition », *op. cit.*
- 29 Le philosophe Michael Polanyi décrit par exemple pour cette raison la reconnaissance des visages comme la forme d'un savoir « implicite », c'est-à-dire ni explicitable ni formalisable, voir Michael Polanyi, *Implizites Wissen*, Francfort-sur-le-Main : Suhrkamp 1985 [1966], p. 14 et suiv., p. 20, p. 27.
- 30 Voir par ex. Hadyn Ellis, « Theoretical Aspects of Face Recognition », in : Graham Davies et al., *Perceiving and Remembering Faces*, Londres et al. : Academic Press 1981, p. 171-196.
- 31 De la même manière, il n'est pas une seule publication de psychologie cognitive de la perception des visages qui puisse faire l'économie de la prise en compte de l'état de la reconnaissance automatisée de formes. Voir par ex. Vicki Bruce, *Recognising Faces*, Hove et al. : Routledge 1988.
- 32 Gilles Deleuze, Félix Guattari, *Mille plateaux*, *op. cit.*, p. 206, 217, 233.
- 33 Voir Matteo Pasquinelli, « Der italienische Operaismo und die Informationsmaschine », in : Ramón Reichert (éd.), *Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*, Bielefeld : transcript 2014, p. 313-332, ici : p. 321.

- 34 Gilles Deleuze, Félix Guattari, *Mille plateaux*, *op. cit.*, notamment p. 216-219.
- 35 *Ibid.*, p. 221.
- 36 Matthew Turk, Alex Pentland, « Eigenfaces for Recognition », in : *Journal of Cognitive Neuroscience* 3.1 (1991), p. 71-86, ici : p. 71.
- 37 Au sujet des principes de base de ce procédé voir L. Sirovich, M. Kirby, « Low-Dimensional Procedure for the Characterization of Human Faces », in : *Journal of the Optical Society of America* 4 (1987), p. 519-524.
- 38 Voir Lila Lee-Morrison, *Portraits of Automated Facial Recognition. On Machinic Ways of Seeing the Face*, Bielfeld : transcript 2019, p. 56 et suiv.
- 39 Turk, Pentland, « Eigenfaces for Recognition », *op. cit.*, p. 83.
- 40 Voir Lee-Morrison, *Portraits of Automated Facial Recognition*, *op. cit.*, p. 79.
- 41 Voir l'approche plus détaillée du sujet dans Meyer, *Operative Porträts*, *op. cit.*, p. 337-356.
- 42 Voir à ce sujet Roland Meyer, « Lichtbildbelehrungen. Bilder im Grenzbereich. Die e-Pass-Fotomuster-tafeln der Bundesdruckerei », in *Bildwelten des Wissens* 4.2 (2006), p. 64-68.
- 43 Voir Ethem Alpaydin, *Machine Learning : The New AI*, Cambridge (Mass.) : The MIT Press 2016, p. 24 et suiv., p. 65 et suiv.
- 44 Voir Gary Huang et al., « Labeled Faces in the Wild : A Database for Studying Face Recognition in Unconstrained Environments » (2008), <http://vis-www.cs.umass.edu/papers/lfw.pdf> [dernier accès 10/03/2020].
- 45 Voir Rainer Mühlhoff, « Menschengestützte Künstliche Intelligenz. Über die soziotechnischen Voraussetzungen von "deep learning" », in : *Zeitschrift für Medienwissenschaft* année 11 (2019), cahier n° 21, p. 56-64, ici : p. 62.
- 46 Yaniv Taigman et al., « DeepFace. Closing the Gap to Human. Level Performance in Face Verification », in : *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus : 2014, p. 1701-1708, <https://ieeexplore.ieee.org/document/6909616> [dernier accès 08/06/2020].
- 47 Je dois cette expression à Simon Rothöhler. Voir à ce sujet et plus généralement à propos de la logistique des flux d'images numériques également *idem.*, *Das verteilte Bild. Stream. Archiv. Ambiente*, Munich : Fink Verlag 2018.