



# TEXTO PARA DISCUSSÃO

ISSN 0103-9466

461

**O enigma de Paul Virilio:  
visão maquínica e metamorfose da percepção**

**Eduardo Barros Mariutti**

**Abril 2024**



**ie** Instituto de  
economia

# O enigma de Paul Virilio: visão maquínica e metamorfose da percepção

Eduardo Barros Mariutti \*

## Resumo

Este é um artigo exploratório que parte da indagação de Paul Virilio sobre a natureza e as características das *representações internas* que as “máquinas que vêem” produzem ao gerarem uma *percepção automatizada e multispectral* da realidade. Escrevendo no final de década de 1980, ele antecipou muito do que viria a ocorrer muito tempo depois: a constituição de uma interpretação *preditiva e estatística* da realidade cada vez mais permeada nas instituições disciplinares e na conduta da guerra. Contudo, ele exagerou ao sugerir que estas representações internas das máquinas, por não precisarem de nenhuma saída videográfica, tenderiam a excluir o homem do *imaginário mecânico* que passaria a agenciá-las. Para contornar essa visão escatológica, argumento que a noção de *visão maquínica* proposta por John Johnston é muito mais adequada: explicita as diferenças entre a percepção automatizada das máquinas e a humana, mas explora o terreno comum que liga ambas.

**Palavras-chave:** Paul Virilio, 1932-2018, Automação da percepção, Visão maquínica.

## Abstract

### *Paul Virilio's enigma: machinic vision and the metamorphosis of perception*

This is an exploratory article that stems from Paul Virilio's inquiry into the nature and characteristics of the internal representations that "vision machines" produce by generating an automated and multi-spectral perception of reality. Writing at the end of the 1980s, he anticipated much of what would happen much later: the constitution of a predictive and statistical interpretation of reality increasingly permeated by disciplinary institutions and the conduct of war. However, he exaggerated when he suggested that these internal representations of machines, because they didn't need any videographic output, would tend to exclude man from the mechanical imaginary that would come to act on them. To circumvent this eschatological view, I argue that the notion of machinic vision proposed by John Johnston is much more appropriate: it makes explicit the differences between the automated perception of machines and that of humans but explores the common ground that links the two.

**Keywords:** Paul Virilio, 1932-2018, Automation of perception, Machinic vision.

**JEL:** A19.

Nós, na verdade, nos revezamos entre estados de relações sociais e estados de relações não-humanas, mas isso não é o mesmo que nos revezarmos entre humanidade e objetividade. O equívoco do paradigma dualista foi sua definição de humanidade. Até a forma dos humanos, nosso próprio corpo, é composta em grande medida de negociações e artefatos sociotécnicos. Conceber humanidade e tecnologia como polos opostos é, com efeito, descartar a humanidade: somos animais sociotécnicos e toda interação humana é sociotécnica. Jamais estamos limitados a vínculos sociais. Jamais nos defrontamos unicamente com objetos (Bruno Latour, 2001, p. 245).

---

\* Professor Associado do Instituto de Economia da Unicamp e do Programa de Pós-Graduação *San Tiago Dantas*. Membro da Rede de Pesquisa em Autonomia Estratégica, Tecnologia e Defesa (PAET&D). E-mail: [mariutti@unicamp.br](mailto:mariutti@unicamp.br). ORCID: <https://orcid.org/0000-0002-3674-3194>.

## Introdução

Paul Virilio abre o capítulo 5 de *La Machine de Vision* com a frase “Agora os objetos me percebem” alegadamente proferida pelo pintor Paul Klee na década de 1920 e, na sequência, afirma que o que parecia ser uma expressão surpreendente – quase uma *boutade* – recentemente (1988) se tornou um fato objetivo, uma realidade: câmeras e sensores controlados por computadores sem a intervenção humana direta estão conseguindo não apenas reconhecer o contorno e a forma das figuras, mas também *interpretar* o que acontece ao seu redor. Este processo de *automação da percepção* evidencia uma questão fundamental: como operam as *representações internas* da realidade nas máquinas perceptivas? Sabemos que se trata de uma apreensão matemática e estatística, isto é, no interior de cada máquina formam-se *imagens instrumentais virtuais* que podem ser diretamente comunicadas para outras máquinas, imagens que não precisam de um suporte ou uma saída videográfica. Precisamente por conta disto, do nosso ponto de vista, elas representam *um enigma* que não temos como decifrar, pelo menos não nos termos dos nossos sistemas internos de representação e de formação de imagens.

Trata-se de uma indagação bastante sagaz, especialmente se levarmos em conta o momento em que foi proferida, quando a computação digital ainda engatinhava, dadas as limitações de memória e de capacidade de processamento dos computadores mais avançados da época. Este artigo exploratório parte desta questão e tenta apontar algumas possibilidades que não foram explicitadas por Virilio. O principal limite da sua interpretação é a ideia de que esse *imaginário* que as máquinas começaram a criar com o processo de automação da percepção tenderia a se autonomizar completamente e, no limite, a excluir o homem de seus agenciamentos. Creio que a noção de *visão máquina* proposta por John Johnston é muito mais adequada: explicita as diferenças entre a percepção automatizada das máquinas e a humana, mas explora o terreno comum que liga ambas. A fronteira existe, mas ela é cada vez mais porosa, principalmente se levarmos em conta as novas interfaces homem-máquina desenvolvidas nos últimos 30 anos, que utilizam sons, vibrações e tecnologias hápticas que criam um grau de integração e de experiências imersivas inimagináveis no final da década de 1980, onde a tela, o teclado e o mouse eram praticamente a única forma de interação do homem com os computadores.

## Automação da percepção e o enigma de Paul Virilio

O final da década de 1980 foi marcado pela proliferação de câmeras de vigilância no interior dos estabelecimentos comerciais e no ambiente urbano nos países desenvolvidos. Isto se deu em meio a transformações significativas no plano da cultura visual, cada vez mais impactada pela arte digital e seus modelos de visualização e de fabricação de imagens. Os modelos analógicos de transmissão de imagens então hegemônicos até meados da década de 1970 – televisão, fotografia e cinema – eram miméticos, isto é, correspondiam aos comprimentos de onda do espectro visual e a um *ponto de vista* situado em um “espaço real” (Crary, 1992, p. 1-2). Neste caso, a relação entre *observador* e *representação* permanecia congruente com o *perspectivismo cartesiano*, isto é, um regime escópico baseado em uma ordem espaço-temporal isotrópica, ocupada por objetos discretos e internamente estáveis percebidos por um *ponto de vista* focado e incorpóreo. (Jay, 2020). Isso começou a mudar com a digitalização do mundo analógico e com a produção direta de *objetos digitais* viabilizada por computadores capazes de gerar imagens eletrônicas e transmiti-las instantaneamente. Ao serem construídas por dados, as imagens se tornam fabricadas e fungíveis, dissolvendo o mundo oticamente

estruturado que marcou a modernidade (Crary, 1992, p. 2). A relação entre observador e representação muda radicalmente, tanto por conta da dissolução do observador quanto pelo fato das imagens poderem ser construídas sem nenhuma relação direta com o “real”.

Paul Virilio foi um dos primeiros a perceber esta dinâmica. Em 1988 ele chamou a atenção para o fenômeno da *automação da percepção* subjacente às *máquinas dotadas de visão*, isto é, dispositivos que, por meio de sensores comandados por computadores, são capazes de interpretar o campo visual no seu entorno, produzindo uma *visão sem olhar* (*vision sans regard*): uma percepção  *sintética* do real muito peculiar, construída matematicamente pela *comunicação* entre máquinas baseadas em códigos informacionais que não necessitam de uma saída videográfica ou qualquer outro tipo de anteparo – o quadro, a gravura etc. – que possibilite a visualização e o entendimento humano (Virilio, 1988, p. 125). Outro aspecto importante é que, muito precocemente, Virilio percebeu é o caráter  *preditivo* da interpretação automatizada da realidade. O *tempo* das percepções sintéticas não é o dos acontecimentos, mas o de um feixe de possibilidades que se projeta do *atual* e que, por conta da suposta insuficiência da profundidade temporal da cognição humana, está além da nossa capacidade de apreensão. O ponto a ser destacado é que essa alegada inadequação não deriva da acuidade do nosso sensorio (algo que pode ser corrigido por próteses sensoriais como telescópios, microscópios, estetoscópio etc.), mas da limitação *temporal* da mente humana em capturar *imagens* e virtualidades de forma consciente. Portanto, a percepção sintética da realidade opera em um espaço-tempo tão veloz e abrangente que é, a princípio, inacessível ao homem, pois a máquina de visão sistematicamente *projeta o futuro imediato como um quadro de possibilidades*, engendrando uma percepção *estatística* da realidade que opera em uma escala temporal diversa da nossa percepção.

Virilio destaca que o ato de ver não é estático e atemporal, pois a visão implica uma *percepção situacional* que é engendrada pela combinação entre memorização e previsão: é isso que confere sentido ao que se vê. Algo aparentemente tão prosaico como atravessar uma rua movimentada ilustra esse argumento. Para fazê-lo com segurança é necessário identificar e memorizar todos os elementos relevantes (automóveis, pedestres, eventuais barreiras etc.), estimar suas posições futuras para poder atravessar no momento correto. Assim como os humanos, sistemas automatizados de vigilância também precisam visualizar e “entender” o que acontece ao seu redor, habilidade que pressupõe alguma forma de percepção da temporalidade, do espaço e dos objetos contidos dentro do seu campo de “visão”. Contudo, como já indicado, essa percepção, além de ser matemática e estatística, pode ser mediada por sensores multiespectrais, que reconhecem elementos e padrões invisíveis ao sensorio humano. Logo, as “imagens” geradas por máquinas e para máquinas assumem uma forma e um significado totalmente independente dos sentidos humanos e da nossa forma de inteligência, ensejando um imaginário mecânico (*imaginaire machinique*)<sup>1</sup> do qual nós, enquanto homens, estamos excluídos (Virilio, 1988, p. 126).

Paul Virilio estava na trilha correta quando afirmou que a *conduta da guerra* estava sendo progressivamente encapsulada e gerida pelo que ele chamou de *logística da percepção*. Em *Guerra e Cinema* (2005) ele afirma que atividade militar sempre envolveu alguma forma de gestão da imagem

---

(1) Esta tradução é controversa. Estou seguindo o entendimento da tradução para a edição em inglês, que grafa *mechanized imaginary*, uma expressão que me parece mais precisa do que *maquínico*, especialmente se levarmos em conta o modo como Deleuze e Guattari utilizam esse termo. A passagem em que a expressão é usada é esta: “*En effet, sans sorties graphique ou vidéographique, la prothèse de perception automatique fonctionnera comme une sorte d’imaginaire machinique dont nous serons cette fois totalement exclus*” (Virilio, 1988, p. 126).

e da comunicação. Mas a Primeira Guerra Mundial acirrou a importância de desenvolver uma *logística da percepção*, isto é, a necessidade imperiosa de criar formas de *visualização* do conjunto das forças em antagonismo no campo de batalha e, também, de se aprimorar sistemas de processamento e de transmissão *de informações* que possibilitem a coordenação e a execução da atividade militar de forma cada vez mais acelerada e *preditiva*. Logo, a logística da percepção desloca o eixo estratégico do campo de batalha não apenas para o conjunto de organização social (como indica a ideia de *guerra total* e, também, as táticas de controle da população usadas pelas forças policiais), mas também para o campo multidimensional da visão e da percepção, o qual é operado por um intrincado conjunto de sistemas sociotécnicos baseados em inteligência artificial e na automação da percepção.

O fato é que um novo *campo* é aberto para as relações de poder em geral e para o conflito militar em particular: o *tempo* ganha preeminência sobre o espaço (Der Derian, 1990, p. 297), assim como a disputa se desloca para a dimensão da *virtualidade*, isto é, para o *feixe de possibilidades* que emergem do atual. Em *La Machine de Vision* Paul Virilio interpreta de forma bastante peculiar deste deslocamento. A ideia de espaço começou a se transformar quando as novas tecnologias da informação viabilizaram a *telepresença*: utilizando dispositivos de comunicação instantânea uma pessoa pode ocupar mais de um espaço ao mesmo tempo e, reversivamente, estes outros “espaços” se manifestam concomitantemente no mesmo campo de percepção. Logo, isto favorece uma desconexão entre os espaços e a presença física das pessoas e suas interações sociais. A telepresença também opera em conexão com a já mencionada partilha da perspectiva entre o homem e as máquinas perceptivas, pois os sensores multiespectrais possibilitam a virtualização de corpos e objetos, multiplicando assim suas representações. Neste caso a imagem em *tempo real* (*temps réel*) – um míssil na tela de um radar, a identificação de um padrão de risco em um metrô vigiado eletronicamente etc. - passa a *dominar* o objeto representado que, na prática, implica a dominação do *virtual* – o feixe de possibilidades – sobre o *atual*, perturbando com isso não apenas o espaço, mas a própria noção de *realidade*<sup>2</sup>. Neste caso, a estrutura do tempo se modifica, dado que o passado, presente e o futuro são substituídos apenas por *dois tempos*, o tempo real (*temps réel*), cada vez mais comandado e acelerado pelo imaginário mecânico, e o *delay* (*temps différé*) (1988, p. 140). O *tempo real* contém sempre o presente e um pouco do futuro em uma escala cada vez mais inacessível ao olho humano e à nossa capacidade de inteligência<sup>3</sup>.

Este é um ótimo ponto de partida para a discussão. Mas é importante matizar esse diagnóstico, fazendo algumas ressalvas. A primeira delas que, a meu ver, contribui para acentuar o catastrofismo da perspectiva de Paul Virilio, é a não observância da *positividade* de alguns desdobramentos teóricos e práticos derivados desta situação. Ele afirma corretamente que o imaginário das máquinas

---

(2) Ele chama esse fenômeno de *lógica paradoxal*, a sucedânea da *lógica dialética* ligada à emergência da fotografia e do cinema no século XIX (i.é., a revelação das dimensões da realidade que não são diretamente acessíveis ao olho humano). “*Le paradoxe logique, c’est finalement celui de cette image en temps réel qui domine la chose représentée, ce temps qui l’emporte désormais sur l’espace réel. Cette virtualité qui domine l’actualité, bouleversant la notion même de réalité. D’où cette crise des représentations publiques traditionnelles (graphiques, photographiques, cinématographiques ...) au profit d’une présentation, d’une présence paradoxale. télé-présence à distance de l’objet ou de l’être qui supplée son existence même, ici et maintenant*” (Virilio, 1988, p. 134).

(3) Quando apresenta esse argumento ele usa como metáfora um míssil rastreado em tempo real: quando ele é capturado pelo radar, o presente mediado pela tela contém o futuro, isto é, o alvo visado pelo míssil (Virilio, 1988, p. 140).

perceptivas é essencialmente probabilístico. Mas ele também sugere que a tendência aparentemente inevitável é a captura de toda a virtualidade por meio dos procedimentos instrumentais construídos e mobilizados por máquinas e para as máquinas, *excluindo o homem* e, ao mesmo tempo, capturando instrumentalmente o futuro. Esta aceleração, além de precipitar o fim do espaço, eliminaria qualquer experiência *qualitativa* do tempo, isto é, o tempo como *duração* tal como o concebem, por exemplo, Fernand Braudel ou Henri Bergson. Seguindo Virilio, poderíamos dizer que o sonho da cibernética – atravessar o animado e o inanimado, o orgânico e o mecânico por laços de informação que *resistem à entropia* – está prestes a se realizar como um pesadelo: toda a vida seria aprisionada por um mecanismo de feedback negativo operando na instantaneidade da comunicação entre as máquinas.

Trata-se de um diagnóstico explicitamente escatológico. Paul Virilio converteu-se ao catolicismo em 1950 e, em pelo menos duas ocasiões, afirmou que não se considera um autor *revolucionário*, mas sim *revelacionário*: a mortalidade é a base da consciência humana, e seu papel como profeta-intelectual é revelar as mistificações que a apropriação da técnica pelo militarismo engendra na cultura visual (Armitage; Bishop, 1998, p. 2-5). Em uma entrevista, confrontado com a pergunta se ele corroborava com o anti-humanismo das perspectivas estruturalistas da década de 1950 e 60, ele respondeu que não, pois era um *anarco-cristão* praticante, inspirado em grande medida em Jacques Ellul:

*I am an anarcho-Christian. It sounds quite paradoxical, but to me the definition of man is subsumed, and I quote it often, in a saying by someone I have come to like very much, Hildegarde from Bingen. St Hildegarde wrote, composed music, played the harp, and was many other things at once. The saying is: 'Homo est clausura mirabilium dei': 'Man is the closing point of the marvels of the universe' (i.e. God). Thus, for me, Man is not the centre of the universe, he is the end of the universe, the end of the world. This has nothing to do with ideas like 'transcendental ego' or 'egocentrism'. For me, there is nothing beyond man. Forget about technology, eugenism, robotics, prostheses. Forget also about [Friedrich Nietzsche's] 'Uebermensch' [Superman]. I do not believe these ideas are at all humanist. I think they're far worse. This is a very important point for me, because I am absolutely against this newfangled form of totalitarianism which I call technoscience and its cult. I see there a yet unheard-of eugenics programme, eugenics written very large, far beyond [Sir Francis] Galton's, the idea behind this new brand of eugenicism being to perfect man, to make a better man. Well, there is no such thing as the possibility of 'improving' man, of tinkering man into something better. No way. Never (Virilio, 2001, p. 20).*

Um estranho humanismo tecnofóbico, onde só resta ao homem decaído esperar pelo apocalipse, pois o culto à tecnologia tende apenas a apressar o nosso fim.

A despeito desta visão escatológica, é importante reconhecer que Virilio enfatiza corretamente que as máquinas perceptivas ajudam a precipitar o fim do humanismo renascentista, dado que elas evidenciam a questão filosófica da *partilha da perspectiva* entre o animado (o ser vivo) e o inanimado (a máquina que cria uma percepção sintética e automatizada do real). O problema é quando ele leva a questão à um cenário limite: a crescente autonomização deste imaginário construído pela interação das máquinas de visão poderia excluir o homem do jogo, dado que para elas só interessa a dimensão sensorial humana enquanto dados de entrada, que não necessitam de uma interface de saída compreensível para os homens. As *representações internas* da máquina – suas imagens virtuais

instrumentais – seriam para nós um *enigma*, similar às figurações mentais de um interlocutor estrangeiro (Virilio, 1988, p. 126).

Esta me parece ser uma questão fundamental: como seria o *imaginário* de máquinas com percepção situacional que são capazes de se comunicar com outras máquinas? Esse enigma seria solúvel? A preocupação central de Virilio é com a *velocidade* e o escopo da percepção. Toda captura de imagem (*prise de vue*) que realizamos é *uma tomada temporal (prise de temps)*. Na câmara fotográfica convencional o tempo de exposição se liga às características da lente e a natureza do *suporte*: a qualidade do acetado de celulose e da emulsão fotossensível que registram a imagem no filme. Pouco tempo de abertura do obturador torna a imagem irreconhecível, uma abertura muito prolongada produz uma imagem velada. Já no caso da visão humana, tudo que é mais acelerado do que a capacidade de processamento do córtex occipital não pode ser percebido<sup>4</sup>. As máquinas são muito mais *rápidas* na tomada das imagens e no seu processamento, assim como possuem um escopo perceptivo muito maior: elas captam todo o espectro eletromagnético, calor, vibração e qualquer outro estímulo que possa ser captado pelos sensores.

O próprio Virilio ressalta que falar em *imagem* não é o mais ideal, pois a *interpretação* que a máquina faz da realidade circundante não guarda nenhuma relação com as imagens que formamos em nossa mente. Para um computador, uma imagem, um som ou um movimento não passam de uma série de *impulsos codificados estatisticamente* que, uma vez internalizados, não precisam mais de nenhum retorno visual. Podem ser armazenados, reprocessados e comunicados enquanto informação para outras máquinas. Por isso o enigma. As máquinas percebem *virtualidades* e reconhecem padrões que não conseguimos ver e temos muita dificuldade em imaginar. Contudo, é pouco plausível a ideia de que o imaginário das máquinas perceptivas poderia se autonomizar completamente, a ponto de tornar o homem efêmero. O problema desta profecia reside precisamente na *exclusão da interface com o homem*, “como se a máquina só pudesse dizer de si mesma e, por conseguinte, nos excluiria do seu modo de existência” (Mintz, 2015, p. 105). O ponto mais óbvio é que o homem participa ativamente no treinamento dos *datasets* que viabilizam os algoritmos que reconhecem imagens e nada indica que essa função será totalmente eliminada. Contudo, a questão mais interessante é pensar a sempre cambiante interface homem-máquina da perspectiva de uma *visão maquínica*, o tema da próxima seção.

Outro aspecto curioso é que, embora Virilio destaque o imbricamento entre a técnica, a política, a conduta da guerra e as *metamorfoses da percepção*, neste caso em particular a técnica estranhamente aparece como algo tendencialmente *divorciado* da sociedade e, mais do que isso, parece ser capaz de se assenhorar completamente dela. Em grande parte, isto se deve ao fato de que Virilio não conseguiu romper completamente com o humanismo, como fica claro no trecho de sua entrevista reproduzido acima. Em seu pensamento persiste a ideia de que o corpo é a unidade do homem e, por extensão, a base do humanismo e do princípio da realidade. Em congruência com este princípio, “ele só é capaz de ver a tecnologia como uma prótese alienígena e externa, intrusiva nas

---

(4) “*Si voir c’est prévoir, on comprend mieux pourquoi la prévision devient, depuis peu, une industrie à part entière, avec l’essor de la simulation professionnelle, de l’anticipation organisationnelle, jusqu’à cette venue des « machines de vision » destinées à voir, à prévoir, à notre place, machines de perception synthétique capables de nous supplanter dans certains domaines, certaines opérations ultra-rapides où nos propres capacités visuelles sont insuffisantes du fait de la limitation, non plus de la profondeur de champ de notre système oculaire comme c’était le cas avec le télescope, le microscope, mais du fait de la trop faible profondeur de temps de notre prise de vue physiologique*” (Virilio, 1988, p. 129).

capacidades naturais do corpo e, conseqüentemente, produzindo apenas efeitos desorientadores e alienantes, tornando-o por sua vez cada vez mais suscetível à manipulação” (Johsnton, 1999, p. 29; 32)<sup>5</sup>.

É importante destacar que Virilio estava reagindo à tendência cada vez mais nítida de *descartar* o papel do corpo na percepção da realidade e na própria existência. Arthur e Marilouise Kroker definiram a década de 1990 como “the flesh-eating 90’s”, um período em que a informação se “desencarna”, isto é, se autonomiza frente aos seus substratos e as tecnologias digitais amparadas no capitalismo desenfreado passaram a “devorar a carne dos homens” (1996), abrindo caminho para utopias *high tech* ao estilo de Hans Moravec (1988), que defendia a possibilidade de se fazer um *upload* da mente em um computador e, deste modo, viver eternamente. No entanto, Virilio exagera demais no peso do corpo enquanto uma unidade discreta, sujeito à pressão “externa” da tecnologia. Basta confrontar a posição dele com a de Donna Haraway (1991), por exemplo, que faz uma crítica e uma reivindicação muito similar, sem ceder ao humanismo e à tecnofobia. Ela tem em comum com Virilio o foco na dimensão *militar* das novas tecnologias ligadas à comunicação e à metamorfose da percepção. Em sua visão, o ciborgue é fruto dos mecanismos de retroalimentação estabelecidos entre o capital e a guerra. Ela também rejeita uma perspectiva centrada no controle informacional dos compostos bióticos, que é compatível com a tensão explicitada por Virilio entre o modo como as máquinas produzem uma visão sintética instrumental e a “percepção humana”. Mas, ao contrário do revelacionário Virilio, Haraway tem um apreço peculiar pela blasfêmia<sup>6</sup> que transparece na sua luta acirrada contra o imaginário e as políticas centradas em um *corpo orgânico* capaz de organizar a resistência contra a tecnologia pensada como uma *exterioridade*, a qual deriva dos dualismos que ainda persistem (mente e corpo, animal e máquina, idealismo e materialismo etc.) no debate público. Ela *celebra* o ciborgue – “um híbrido de máquina e organismo, uma criatura de realidade social e também uma criatura de ficção” – pois ele subverte a relação entre a natureza e a cultura, possibilitando pensar essa questão em uma outra chave: por ser um símbolo da dominação tecnológica absoluta sobre a existência e ao mesmo tempo uma metáfora da empatia e parentesco (*kinship*) entre homens, animais e máquinas, o ciborgue explicita tanto as formas de dominação quanto as possibilidades *inimagináveis de outro ponto de vista*. Aqui a ruptura com a dimensão escatológica do pensamento de Virilio é explícita.

---

(5) Quem tem intimidade com a obra de Virilio tende a concordar com a síntese proposta por Johnson sobre esse assunto: o longo período da *percepção natural* (i.é., ancorada no corpo e no sensorio humano) é encerrado abruptamente com a invenção do telescópio e com a matematização da natureza e do campo visual. As próteses de visão ajudam a corroer a fé na “visão natural”. Há uma realidade *além* do olho nu. Contudo, o segundo aspecto parece ser mais importante. A ideia de que a lógica matemática da natureza só pode ser apreendida por novos instrumentos que produzem uma leitura diferente da nossa experiência sensorial direta *muda* a ideia de verdade: ela escapa da alçada da visão humana, desancorando o real das capacidades naturais do corpo humano. É por conta disto que ele conclui que a criação das máquinas perceptivas automatizadas irá inevitavelmente tornar a visão humana descartável (cf. Johsnton, 1999, p. 30-32).

(6) Haraway foi educada em uma tradição católica: “*I am conscious of the odd perspective provided by my historical position a PhD in biology for an Irish Catholic girl was made possible by Spunik’s impact on US national science-education policy. I have a body and mind as much constructed by the post-Second World War arms race and cold war as by the women’s movements*” (Haraway, 1990, p. 173). Talvez isso ajude a explicar o seu gosto pela blasfêmia, pois isso exige levar as coisas um pouco a sério, ao contrário da apostasia.

### **Visão maquinica e transformações ontológicas: o espaço vetorial multidimensional.**

Dialogando criticamente com a démarche de Virilio, John Johnston propõe uma via que parece mais frutífera, pois ele refuta as interpretações baseadas na *oposição* entre o homem e a técnica para explorar *o terreno comum* que os conecta, utilizando para isso a expressão *visão maquinica*, que ele define desse modo:

*Machinic vision, as I shall use the term, presupposes not only an environment of interacting machines and human-machine systems but a field of decoded perceptions that, whether or not produced by or issuing from these machines, assume their full intelligibility only in relation to them (Johnston, 1999, p. 27).*

Nesta abordagem a questão se transfigura radicalmente. O pressuposto básico é a presença significativa de interações máquina-máquina (m2m) e homem-máquina que cria um campo de percepções *decodificadas* que só fazem sentido se referidas a esse ambiente. Logo, embora possivelmente repletas de tensão e de pontos de descontinuidade, a visão maquinica engloba um campo de percepções compartilhadas entre homens, instituições e sistemas de máquinas.

Um pouco mais a frente, em outra passagem, ao discutir as implicações da generalização das imagens eletrônicas e sua relação com o cinema na obra de Deleuze, o argumento fica ainda mais claro:

*The digital image is a decoding because it frees the cinematic image from its material support, mobilizing it within a communicational network wherein it can be transmitted anywhere instantaneously; but it is also a recoding because, instead of being inscribed directly on a chemically treated surface, light is converted into information, mathematical data whose infinitesimal discreteness allows the real to be synthesized or recomposed (Johnston, 1999, p. 39).*

A produção da imagem digital é decodificadora porque *liberta* a imagem do substrato e, ao mesmo tempo, ao ser convertida em *informação*, isto é, em dados matemáticos estruturados e plenamente manipuláveis, ela é também recodificadora. Em outros termos: a imagem se torna uma das múltiplas formas que a informação, uma vez abstraída do seu substrato original, pode assumir. A proposta de Johnston é, portanto, explorar esse campo de *percepções* codificadas e recodificadas que pressupõe um imbricamento efetivo entre o orgânico e o eletromecânico, no qual as fronteiras entre homem e máquina – ou, se preferir, do humano e do não-humano – são reiteradamente dissolvidas e reconstituídas sob novas formas.

O fato é que a imagem digital se autonomizou. Ela deixou de ser apenas um ícone ou um simulacro de algo que existente no mundo, isto é, versões digitais de objetos ou entidades localizadas no “mundo real”, tal como era o caso dos processos de digitalização instituídos no final do século XX. Hoje, por meio dos agenciamentos maquinicos, é possível fabricar novas realidades, fato que ajuda a moldar e a tensionar a percepção humana e o “imaginário” das máquinas perceptivas. Como já foi sugerido, o plano em que essas imagens se instanciam não é mais o do olho humano e das próteses analógicas da visão: é um plano *multiespectral*. Só podemos “ver” os estímulos imperceptíveis pelo sensorio humano por meio dessas máquinas perceptivas e seus sensores, que precisam gerar alguma forma de visualização que faça sentido para nós. Contudo, há zonas e

elementos da realidade que as máquinas têm muita dificuldade de apreender. O sistema CAPTCHA<sup>7</sup> destinado a impedir que *bots* se passem por humanos na internet é o exemplo mais óbvio da limitação da visão computacional (Mintz, 2015, p. 146), mas técnicas de ofuscamento das câmeras de reconhecimento facial e os ataques adversariais estão se tornando cada vez mais frequentes (Pasquinelli; Joller, 2020, p. 17-18)<sup>8</sup>.

Logo, persiste uma espécie de *fronteira* entre o que chamamos de visão humana e a visão computacional. Por mais que os sensores multiespectrais acessem zonas que não produzem ressonância em nosso sensorio, as máquinas não podem ver tudo:

Numa espécie de contraponto a tal onividência, contudo, um dos traços mais evidentes dos agentes computacionais de visão talvez sejam justamente os dispositivos desenvolvidos visando seus limites: os CAPTCHA. Diante de uma figura ruidosa de letras e números distorcidos, somos solicitados a identificar os caracteres para seguir adiante na navegação. Enquanto um teste cujo objetivo é distinguir, entre aqueles que acessam determinada página na web, os humanos dos *bots* (literalmente, robôs), os CAPTCHA jogam com uma espécie de fronteira da visibilidade computacional: visíveis para humanos, mas não para máquinas. A necessidade atual de distinguirmo-nos de agentes computacionais através de um tal dispositivo não apenas revela a expectativa de certo nível de desempenho destes programas e de sua presença e atuação ubíqua na web como também parece expressar uma ambiguidade do desenvolvimento de tal tecnologia, dado que em determinadas circunstâncias julgamos necessário recorrer à possibilidade de ainda demarcar a diferença entre visão humana e visão da máquina. A falha dos dispositivos de visão contemporâneos, portanto, longe de um não dito desta tecnologia, integra a dinâmica de seu desenvolvimento, apontando para a complexidade do trânsito de fronteiras entre o humano e a máquina que ela parece inicialmente sugerir (Mintz, 2015, p. 26).

É precisamente a existência dessa fronteira que explicita a importância e a peculiaridade da noção de *visão maquínica*, enquanto um tipo de visão que não se reduz à visão computacional e, ao mesmo tempo, suplementa e transforma as nossas formas de visualidade e de subjetividade.

---

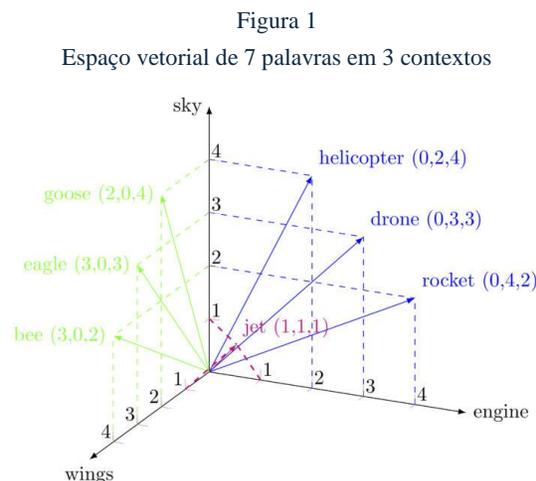
(7) Trata-se de um acrônimo da expressão “*Completely Automated Public Turing test to tell Computers and Humans Apart*”, que é particularmente reveladora: uma variante simples do teste de Turing se tornou necessária com a migração massiva de serviços públicos, operações bancárias, práticas comerciais e profissionais para a internet.

(8) Como muitos sistemas de identificação facial precisam usar a luz infravermelha para capturar imagens, basta usar emissores desta luz para “cegar” os dispositivos (técnica chamada de *inundação*). Existem óculos especiais que refletem a luz de uma forma que confunde os sensores. Estão surgindo também vestimentas com padrões de estampa que confundem os algoritmos, parecendo que existem milhares de rostos onde só há um. Já os ataques adversariais envolvem adicionar pequenos detalhes, geralmente imperceptíveis ao olho humano, que confundem os sistemas de classificação de imagens. Um adesivo colado em uma placa de Pare pode confundir um veículo autônomo, fazendo-o classificar o sinal como uma placa com outro conteúdo, podendo causar acidentes (Papernot; McDaniel; Swami et al., 2017, p. 1-2). Esta tática faz parte do que se convencionou chamar de exemplos (ou ataques) adversariais, que exploram os pontos cegos e as anomalias dos algoritmos que classificam imagens, produzindo falsos resultados. “Adversarial attacks exploit blind spots and weak regions in the statistical model of a neural network, usually to fool a classifier and make it perceive something that is not there. In object recognition, an adversarial example can be a doctored image of a turtle, which looks innocuous to a human eye but gets misclassified by a neural network as a rifle. Adversarial examples can be realised as 3D objects and even stickers for road signs that can misguide self-driving cars (which may read a speed limit of 120 km/h where it is actually 50 km/h). 63 Adversarial examples are designed knowing what a machine has never seen before. This effect is achieved also by reverse-engineering the statistical model or by polluting the training dataset” (Pasquinelli; Joller, 2020, p. 18).

Estamos, contudo, testemunhando o início de uma nova transformação, que foi impulsionada pelo extravasamento das técnicas de *deep learning* orientadas para o reconhecimento de imagens para outras tarefas automatizadas, tais como os *Large Language Models (LLMs)* como o chat GPT, reconhecimento de fala (em celulares, assistentes pessoais, telemarketing etc.), análise de séries temporais (monitoramento de ações no mercado financeiro, meteorologia, previsão de demanda energética etc.), diagnósticos por imagem, detecção de fraudes e muitas outras. Do ponto de vista da computação e da matemática aplicada, qualquer objeto ou padrão pode ser representado digitalmente em um novo tipo de espaço, o espaço vetorial multidimensional. Este espaço é formado pela conversão dos objetos – palavras, sons, imagens etc. – em vetores distribuídos em múltiplas dimensões, ordenados pela sua *vizinhança matemática*. Tomemos um LLM como exemplo. Depois de analisar uma quantidade gigantesca de textos, ele “aprende” que a palavra gato está próxima de cachorro, por exemplo. Assim como marido está próximo de esposa. Cada palavra é convertida em uma *representação numérica* que marca a distância de outras palavras em várias dimensões:

*Words are too complex to represent in only two dimensions, so language models use vector spaces with hundreds or even thousands of dimensions. The human mind can't envision a space with that many dimensions, but computers are perfectly capable of reasoning about them and producing useful results (Lee; Trott, 2023).*

Para se ter uma noção, a palavra Cat é representada em alguns modelos com 300 vetores<sup>9</sup>. Como as representações são numéricas e relacionadas com a proximidade semântica das palavras, é possível fazer operações matemáticas e identificar correlações inusitadas. A Figura 1 retirada [daqui](#) é um exemplo extremamente simplificado destas representações vetoriais:



(9) Esses são os primeiros e os três últimos vetores da palavra Cat de um modelo alimentado pela Wikipedia em inglês: 0.007398007903248072, 0.0029612560756504536, -0.010482859797775745, 0.0741681158542633, 0.07646718621253967, -0.0011427050922065973, (...) -0.0929502621293068, 0.05977592244744301, 0.00015643733786419034. Para ver a sequência completa, acesse: [http://vectors.nlpl.eu/explore/embeddings/en/MOD\\_enwiki\\_upos\\_skipgram\\_300\\_2\\_2021/cat\\_NOUN/](http://vectors.nlpl.eu/explore/embeddings/en/MOD_enwiki_upos_skipgram_300_2_2021/cat_NOUN/).

Como os algoritmos operam com milhares de dimensões e de parâmetros, somos incapazes de compreender o modo como os computadores produzem os seus resultados e, principalmente, como eles formam suas percepções da realidade.

Uma vez convertidos em vetores, a despeito de suas diferenças, os objetos digitais tornam-se entidades *comensuráveis* por meio das suas relações de distância matemática. Como salienta com muita precisão Nuria Rodriguez-Ortega (2022, p. 3), não há nenhum correlato empírico para o espaço vetorial multidimensional. Trata-se de uma *generalização intelectual* fundamentada na possibilidade lógica de representar vetorialmente objetos em múltiplas dimensões, estabelecendo com isso correspondências inusitadas e, portanto, inimagináveis a partir de uma perspectiva tridimensional como a nossa. Seu argumento é que essa prática produz transformações *ontológicas*:

*The current relevance of the concept of n-dimensional space for cultural analysis and interpretation must be related, in the first place, to the «ontological» transformation that cultural objects have been undergoing for decades as a result of their digitization or direct digital production: cultural objects, regardless of their nature (images, words, sounds, etc.), in their digital mode of existence are essentially matrices or sets of numerical data. It is precisely this transformation that makes their computation possible. From a computational point of view, images or texts are nothing more than a spatial surface of numerical information from which it is possible to extract (also) numerical characteristics using computational systems. Therefore, this ontological transformation also implies an epistemological transformation insofar as cultural objects that are transformed into digital forms become a problem of a computational and mathematical order (Rodriguez-Ortega, 2022, p. 2).*

Logo, qualquer objeto pode ser convertido em uma matriz numérica e instanciado em um espaço topológico com múltiplas dimensões que consiste em uma espécie de *taquigrafia* do conjunto dos dados e de suas relações (i.e., um *modo digital de existência*). Por envolver uma representação compacta de variáveis e características não observadas de forma direta, é comum também usar a expressão *espaço latente*, pois ele engloba padrões e propriedades emergentes que não podem ser inferidos diretamente dos dados.

Esse aspecto é decisivo. Cada objeto ocupa uma posição em uma *rede de relações* baseada em sua distância matemática com relação a todos os demais<sup>10</sup>. Essa espacialidade não é fixada por coordenadas externas como a geometria euclidiana, pois é organizada a partir de si própria, isto é, possui apenas dimensões *intrínsecas* (Lury; Parisi; Terranova, 2012, p. 7; DeLanda, 2002, p. 11-13). Esta forma de espacialidade consiste no princípio básico das *representações internas* dos sistemas de aprendizado de máquina:

*Despite the variety of machine learning applications, all machine logic is based on the same set of operations: taking the set of instances or data of the training dataset as a base, machine learning technologies recognize and extract the statistical distribution of patterns that define the (non-directly observed) dataset structure. This statistical distribution of patterns constitutes a*

---

(10) “Consider, for example, the concept of distance, which is a key concept in cultural interpretation that has been the subject of different conceptualizations and definitions over time. In a vector space, distance becomes a mathematical distance. This issue is well illustrated by the so-called cosine similarity, which is one of the basic operations taken from linear algebra that consists of calculating the similarity between two vectors by taking, as a measure, the cosine of the angle between them” (Rodriguez-Ortega, 2022, p. 11).

*statistical model that stands as a representation of the corpus (or domain of reality) under analysis. (...) However, what I want to highlight for now is the condition of latent spaces as models that encode a representation of the cultural domain under analysis as a whole, not only of an individual element. It can be said that this space, which is based on a quantitative spatial representation, is the way a computational system understands the entire domain (Rodríguez-Ortega, 2022, p. 6).*

É essa forma de entendimento da estruturação de todo o domínio que permanece oculta para nós, pois o que está em jogo não é apenas o elevado número de dimensões, mas também as propriedades emergentes e a racionalidade algorítmica inscrita no espaço latente<sup>11</sup>. Para mitigar este estranhamento são cada vez mais frequentes técnicas de redução de dimensionalidade, isto é, formas de representar em gráficos de duas ou três dimensões as relações multidimensionais dos espaços latentes. É provável que outras formas de tentar explorar a porosidade da fronteira entre as nossas representações e a visão computacional surjam nos próximos anos.

De qualquer modo, a digitalização do mundo e a produção direta de objetos digitais está impondo uma compreensão *topológica* do espaço.

*If the concept of an n-dimensional space implies the problematization and reconsideration of some kind of concept, it is precisely the notion of «space» itself. In this sense, I believe that the toy examples of the previous section, although simple, adequately illustrate the fact that these spaces do not act as containers or supports in the traditional sense of the term - that is, spaces where cultural objects, visual contents or writing signs are located, arranged or situated - but are generated by them once they have been transformed into numerical information. Sign and space constitute an indissoluble unity: space is sign and sign is space, and this lack of distinction shifts us from the traditional concept of Euclidean space - according to which space is a pre-existing category, different from what is located in it (it is another dimension, extrinsic to the object) - to the topological concept of space, according to which space is an a posteriori that is generated by and imbricated in the objects themselves*

A noção de um espaço *isotrópico* e inerte foi fundamental na revolução científica do século XVII e preservou a sua influência até pelo menos o começo do século XX, quando a noção de *espaço-tempo* passou a colocar em questão a concepção clássica de tempo. A generalização das máquinas perceptivas e dos dispositivos de *machine learning* estão promovendo outra transformação radical, favorecendo a consolidação de uma *cultura topológica*, na qual objetos, espaço e signos se fundem e, ao mesmo tempo, passam a ser relacionados principalmente por meio de sua distribuição espacial plotada em um campo multidimensional.

### Considerações finais

A ideia de visão maquínica pressupõe que há uma fronteira entre as representações internas e os modos de ver dos homens e a visão computacional, mas que não é absoluta: a despeito das linhas

---

(11) “As I pointed out earlier, making a mental image of a high-dimensional space is very complicated since it is practically impossible to process using our human brain, which is used to thinking - at least until recently - in four-dimensional parameters. The counterintuitive character of these vector spaces lies not only in the number of dimensions but also in some of their properties and in the algorithmic rationale that operates on and produces them, which are sometimes unintelligible to our forms of understanding” (Rodríguez-Ortega, 2022, p. 7).

de corte, existem zonas de sobreposição e espaços de ressonância. É exatamente na *porosidade* desta fronteira que se situam as principais tensões e, sobretudo, as novas possibilidades que as máquinas perceptivas abrem para a vida social. Hoje somos praticamente forçados a conviver com imagens fabricadas e artefatos digitais que não encontram nenhuma contrapartida no mundo “analógico”. Isso, seguramente, ajuda a transformar a nossa percepção da realidade. Além disto, estamos rodeados por sistemas preditivos automatizados (análises de crédito, proteção de fraude financeira, sugestão de conteúdo, motores de busca, automação residencial etc.) que sobredeterminam nossa vida familiar, profissional e os momentos de lazer. Muitos deles habitam um tipo novo de espaço, o espaço vetorial n-dimensional. O argumento central desse estudo exploratório é que esta convivência está produzindo *transformações ontológicas*. Nosso mundo está mudando, dado que a *visão maquínica* nos permite um acesso multiespectral à realidade e, sobretudo, nos expõe a comensurabilidades até então inimagináveis, as quais só podem ser estabelecidas por meio da representação vetorial de objetos muito diferentes em um espaço ordenado pela vizinhança matemática. É deste espaço com propriedades emergentes que se inferem *padrões* inusitados que estão muito além das possibilidades de cognição humana.

A questão é que, ao contrário do que profetizou Virilio, estes agenciamentos maquínicos estão promovendo novas formas de acoplamento entre o homem e as máquinas, transformando com isso não apenas a nossa vida cotidiana, mas também o nosso entendimento da tessitura e da dinâmica da realidade: a interação corriqueira com a internet, smartphones, gerenciadores de energia domésticos (smart meters), aplicativos de navegação por GPS, dispositivos de segurança controlados por assistentes virtuais, edifícios inteligentes e toda a parafernália agregada no conceito de “casa inteligente” ajudam a expandir o escopo e a variedade das interações entre homens, instituições sociais e sistemas de máquinas. O que parece estar em curso é a criação de novas interfaces que sinalizam e ajudam a impulsionar novas formas de relação homem-máquina que ultrapassam a tensão entre o organicismo e o mecanicismo que marcou o século XIX e a maior parte do século XX, favorecendo o estabelecimento de uma suposta “cultura topológica”. Isso exige um estudo muito mais detalhado. Há bastante bibliografia sobre o assunto, mas o tema ainda é bastante escorregadio. O vínculo mais direto com esse estilo de análise em formação é com o que se convencionou chamar de *Science and Society Studies*, ramo genuinamente interdisciplinar que se constituiu pela erosão da fronteira entre a técnica e a sociedade, recorrendo ao “conceito de uma rede híbrida ou de agenciamentos heterogêneos, compostos de múltiplas maneiras por entidades sociais, técnicas e naturais como a melhor maneira de entender o papel da tecnologia na vida social” (Marres, 2012, p. 289). O ponto decisivo é que, nessa “cultura” o foco se desloca das unidades discretas e dos processos lineares para uma reflexão assentada na multiplicidade e variedade de elementos e, sobretudo, incide sobre as formas ou modos de relação que os conectam estruturalmente em um espaço de possibilidades (DeLanda, 2002, p. 10). Isso ajuda a erodir a “concepção escalar tradicional de espaço”<sup>12</sup> fortemente associada ao perspectivismo cartesiano sem que, contudo, possamos ter alguma clareza sobre como as novas formas de se viver o tempo e o espaço vão efetivamente se concretizar.

---

(12) Os entusiastas da cultura topológica entendem a “concepção tradicional de espaço” como uma espacialidade homogênea e isotrópica, baseada na geometria euclidiana que toma os objetos como entidades internamente estáveis, regulares e posicionadas linearmente no tempo e no espaço. Desse modo, as distâncias, comprimento, áreas e volume sempre são expressas de forma fixa: o espaço é métrico. “*There exist other spaces, however, where fixed distances cannot define proximities since distances do not remain fixed. A topological space, for example, may be stretched without the neighbourhoods which define it changing in nature. To cope with such exotic spaces, mathematicians have devised ways of*

### Referências bibliográficas

- ARMITAGE, John. (Org.). *Virilio live: selected interviews*. London: Sage, 2001.
- ARMITAGE, John; BISHOP, Ryan. (Orgs.). *Virilio and visual culture*. Edinburgh: Edinburgh U. Press, 2013.
- CRARY, Jonathan. *Techniques of the observer: on vision and modernity in the Nineteenth Century*. Cambridge: MIT Press, 1992.
- DeLANDA, Manuel. *Intensive science and virtual philosophy*. London: Continuum, 2002.
- HARAWAY, Donna. *Simians, cyborgs, and women*. New York: Routledge, 1991.
- HAYLES, Nancy K. *How we became Posthuman*. Chicago: Chicago U. Press, 1999.
- JAY, Martin. Regimes escópicos da modernidade. *ARS*, n. 38, 2020.
- JOHNSTON, John. Machinic vision. *Critical Inquiry*, n. 26 (1999).
- KROKER, Arthur; KROKER, Marilouise. *Hacking the future: stories for the flesh-eating 90's*. Montreal: New World Perspectives, 1996
- LURY, Celia; PARISI, Luciana; TERRANOVA, Tiziana. Introduction: the becoming topological of culture. *Theory, Culture & Society*, n. 29, v. 4/5, 2012.
- MARIUTTI, Eduardo. *Ontologias informacionais: notas sobre o pensamento de Luciano Floridi*. Campinas: Unicamp. IE, 2021. (Texto para Discussão, n. 421).
- MARRES, Noortje. On some uses and abuses of topology in the social analysis of technology. *Theory, Culture & Society*, n. 29, v. 4/5, 2012.
- MORAVEC, Hans. *Mind children: the future of robot and human intelligence*. Cambridge: Cambridge U. Press, 1988.
- MINTZ, André Góes. *Visão computacional e visualidades contemporâneas: composições do ver e do visível entre a técnica, a ciência e a arte*. Dissertação (Mestrado)—Belo Horizonte: UFMG, 2015.
- PAPERNOT, Nicolas; McDANIEL, Patrick; GOODFELLOW, Ian; JHA, S.; CELIK, Z. B.; SWAMI, A. *Practical black-box attacks against machine learning*. ASIA CCS'17: Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security, 2017.
- PASQUINELLI, Matteo; JOLER, Vladan. *The nooscope manifested: artificial intelligence as instrument of knowledge extractivism*. KIM research group (Karlsruhe University of Arts and Design) and Share Lab (Novi Sad), 1 May, 2020.
- VIRILIO, Paul. *La machine de vision*. Paris: Galilée, 1988.

---

*defining the property of 'being nearby' in a way that does not presuppose any metric concept, but only nonmetric concepts like 'infinitesimal closeness'.*" (DeLanda, 2002, p. 22). Esse é um ponto importante, que merece um estudo específico, que já estou começando a delinear. Manuel DeLanda já enquadrado bem o problema, mas resta correlacionar a sua visão com as reflexões de Luciano Floridi sobre a infoesfera e o que ele chama de ontologia informacional. Já publiquei aqui um pequeno estudo exploratório sobre o tema (Mariutti, 2021).