HAL open science

L'intelligence artificielle appliquée au patrimoine des techniques manuelles : un dialogue entre anthropologie et ingénierie

Arnaud Dubois, Sotiris Manitsaris

▶ **To cite this version:**

Arnaud Dubois, Sotiris Manitsaris. L'intelligence artificielle appliquée au patrimoine des techniques manuelles : un dialogue entre anthropologie et ingénierie. Cahiers d'histoire du Cnam, 2023, 16 (1), pp.91-104. hal-04356520

HAL Id: hal-04356520

https://hal.science/hal-04356520v1

Submitted on 20 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

L'intelligence artificielle appliquée au patrimoine des techniques manuelles : un dialogue entre anthropologie et ingénierie

Arnaud Dubois
HT2S, Cnam

Sotiris Manitsaris
Directeur-adjoint du Centre de Robotique de l'École des Mines, Université PSL

Au sein du projet de recherche « Mingei : Representation and Preservation of Heritage Crafts » – un programme H2020 (2018-2021) qui explorait les possibilités de représenter et de rendre accessibles, *via* les technologies numériques, les aspects matériels et immatériels de l'artisanat en tant que patrimoine culturel – Sotiris Manitsaris (Mines Paris-PSL) et Arnaud Dubois (HT2S-Cnam) ont collaboré, *via* le consortium, sur la numérisation, la formalisation et la diffusion des gestes techniques du soufflage de verre. À partir de cette expérience commune, ils reviennent dans cet entretien sur les enjeux méthodologique et théorique de l'intelligence artificielle du mouvement, la digitalisation des savoir-faire et l'approche biomécanique du corps au travail.

Arnaud Dubois : ***Peux-tu nous présenter ton travail et tes recherches sur les gestes techniques ?***

Sotiris Manitsaris : Peut-être deux mots sur mon background parce que je n'ai pas de formation d'ingénieur, je suis mathématicien. J'ai fait des mathématiques appliquées pour ma licence. J'ai fait ensuite un master en développement local et dynamiques territoriales. C'est à ce moment que j'ai commencé à m'intéresser aux ressources locales et humaines,

y compris aux compétences, aux savoir-faire locaux et à la façon dont on peut les mettre en valeur avec des nouvelles technologies pour qu'elles puissent potentiellement être converties en ressources économiques. Ensuite, j'ai fait une thèse en intelligence artificielle sur la vision par ordinateur, sur le geste musical, le geste pianistique des doigts. Mon objectif était de capturer les gestes musicaux des doigts et d'apprendre à l'ordinateur à les trier et à les reconnaître afin d'ouvrir un potentiel d'interactions pour de nouvelles interfaces musicales. Depuis la fin de mon master, mon père, qui était professeur à l'université, me disait tout le temps : « *Il va falloir que tu trouves le lien entre ton master et ta thèse !* » D'un point de vue extérieur, cela paraissait en effet déconnecté. J'ai beaucoup brainstormé avec une camarade à moi, une future étudiante du Mastère AIMove[1], qui venait du milieu de la préservation du patrimoine matériel. Finalement le lien que j'ai trouvé était dans la façon d'utiliser ces technologies de capture du mouvement, des mains et des doigts. Il faut savoir qu'à l'époque où j'ai terminé ma thèse en 2010, la capture des doigts c'était très innovant, beaucoup plus que ça l'est aujourd'hui très franchement. L'Unesco avait adopté en octobre 2003 la convention [du Patrimoine Culturel Immatériel[2]] ; c'était une convention avec un impact important, mais comme avec toutes les conventions de l'Unesco, il y a souvent la volonté mais, derrière, les méthodologies et les technologies ne sont pas toujours prises en compte. Mon idée a donc été de proposer cette technologie de capture du mouvement et de l'intelligence artificielle pour la préservation du patrimoine immatériel.

J'ai commencé par étudier la poterie, les gestes des potiers, mais aussi le geste musical, y compris pour les chants rares. J'ouvre une petite parenthèse : quand je dis geste, j'entends aussi mouvement. Dans le chant vocal, il n'y a pas de gestes humains mais il y a les mouvements du conduit vocal, les mouvements des lèvres, les mouvements de la langue, l'ouverture/fermeture des cordes vocales. Tous ces éléments ont donné un premier gros projet européen, le projet « *i-Treasures*[3] », sur lequel on a pu, avec beaucoup d'autres institutions

1 AIMove est un Mastère spécialisé en intelligence artificielle et mouvement pour la robotique et les systèmes interactifs, délivré par Mines Paris – PSL. Ce programme professionnel accorde une grande importance au développement de l'expertise dans les domaines de l'apprentissage automatique et de l'apprentissage profond pour l'interaction et la collaboration homme-machine [URL : https://www.minesparis.psl.eu/Formation/Masteres-Specialises/Masteres-Specialises-temps-plein/AIMOVE/].

2 Convention de 2003 [URL : https://ich.unesco.org/fr/convention].

3 « *i-Treasures* » (Intangible Treasures – Capturing the Intangible Cultural Heritage and Learning the Rare Know-How of Living Human Treasures) est un projet de recherche européen (2013-2017) qui a permis de développer une plateforme ouverte et extensible pour donner accès aux ressources du Patrimoine Culturel Immatériel (PCI) de l'Unesco, et contribuer à la transmission des savoir-faire rares des Trésors Humains Vivants aux apprentis. Sa principale contribution a été la création de nouvelles connaissances en proposant de nouvelles méthodologies et de nouveaux paradigmes technologiques pour l'analyse et la modélisation du PCI.

européennes, travailler sur la transmission du patrimoine immatériel. L'objectif du projet était de proposer une plateforme numérique avec une sorte de boîte à outils qui pourrait être utilisée pour la transmission du savoir-faire. Pour ma part, j'ai travaillé avec des potiers à Vallauris, à Cappadoce en Turquie, et des potiers en Grèce. On a comparé les différentes techniques. J'ai aussi eu l'occasion de travailler sur la conceptualisation de nouvelles interfaces d'interaction homme-machine qui auraient le potentiel de renouveler le patrimoine. C'était une interface appliquée dans le domaine de la musique. Il me semble qu'à l'Unesco, au moment de la Convention, le terme « *transmission* » qui avait été employé n'était pas considéré comme relevant seulement de l'apprentissage humain. La transmission avait aussi un sens d'accès à la connaissance, du fait que le métier de musicien ou ce type de patrimoine existe. En s'appuyant sur ce concept, on a proposé un instrument de musique immatériel, inspiré par le piano (c'est-à-dire inspiré par les gestes musicaux des doigts) mais plus accessible au grand public qu'un piano, surtout pour les enfants, afin qu'ils aient accès à des formes musicales issues des grands compositeurs tels que Beethoven mais avec un geste bien plus facile. On l'appelait à l'époque « *everyday* », de façon à ce qu'il puisse faire un geste et générer de la musique sans avoir forcément à passer par des années et des années d'entraînement sur un piano. Donc il y a eu la préservation mais aussi une dimension relevant du renouvellement. En parallèle, j'ai été recruté aux Mines, au Centre de Robotique, donc il fallait quand même laisser un peu de place pour les robots.

Les robots pour nous n'ont pas pour finalité d'automatiser ou de remplacer l'humain. Tu as entendu parler de la quatrième révolution industrielle ? Les gens disent : « *oui, les robots vont remplacer l'humain* ». L'humain a été remplacé dans les années 1960, 1970. Tout au long de la « troisième révolution industrielle », on avait des robots bêtes et méchants. Donc les opérateurs ont déjà été remplacés en quelque sorte ! Là, on est plutôt sur la complémentarité des compétences entre humains et robots. C'est-à-dire qu'il faut étudier le geste, analyser la dextérité du geste des opérateurs, de façon que l'on puisse, sur cette base-là, entraîner un robot qui puisse venir assister l'humain, en attribuant des tâches pénibles et répétitives aux robots. L'objectif est quand même de libérer la créativité et la dextérité, de désengager l'humain des risques d'avoir des troubles musculosquelettiques à cause des tâches pénibles. Même si, au premier abord, la robotique n'a absolument rien à voir avec les métiers d'art et le patrimoine, je pense que la robotique a un rôle à jouer. Et il faut que nous, les chercheurs, on se prépare pour ça parce que si on ne le fait pas, les robots vont arriver de façon complètement désordonnée et les entreprises et les fabricants de robots vont nous imposer leurs règles du jeu. Il faut donc que les chercheurs, ethnologues, ingénieurs et toutes les parties prenantes soient impliqués, puissent brainstormer

et préparer cette arrivée des robots, même dans les métiers d'art. Nous, on ne fabrique pas la machine. On ne fabrique pas le bras du robot ou un robot mobile. On développe des couches extérieures à la machine. Aujourd'hui, les robots, on les achète. Ce sont des produits d'étagères. Donc on achète un robot et on leur crée des couches extérieures de perception, sur la base des intelligences artificielles qui sont capables de reconnaître le comportement humain et le geste corporel. On part du principe que si le robot a la capacité de reconnaître le geste humain, le geste technique professionnel, il va pouvoir anticiper spatialement et temporellement. Spatialement en prenant en compte le gabarit de l'humain, certains aspects anthropométriques, le fait qu'il soit gaucher ou droitier, et donc passer la pièce à l'humain au bon endroit. Temporellement, en s'adaptant au rythme de l'humain.

Comment définis-tu le robot et l'intelligence artificielle du mouvement ?

Il y a deux grandes écoles. Celle avec la plus grande communauté scientifique et qui met en œuvre ce que je qualifierais d'« approche du robot outil ». Le robot est une extension du corps humain, comme un outil, comme une visseuse, comme un marteau. Il sert uniquement à effectuer une tâche que l'humain aurait plus de mal à mettre en place sans outil, avec ses propres mains. À l'École des Mines, on a une approche plutôt « robot collègue » ou « robot assistant » qui se base sur le principe de la complémentarité des compétences. On ne remplace pas l'humain. On vient l'assister. On parle de transmission des savoir-faire. Est-ce que tout est à transmettre ? Est-ce que les tâches pénibles sont à transmettre ? Ce que l'on propose, c'est une sorte d'analyse du poste de travail, des tâches, une sorte de cartographie du geste et des tâches en fonction de leur pénibilité. Et lorsque l'on souhaite impliquer un robot – parce que l'on n'implique pas toujours un robot – sur la base de cette cartographie, on attribue les tâches pénibles et répétitives au robot. On garde seulement, pour l'humain, les tâches à forte valeur ajoutée. C'est ça le principe. Et tout cela se base sur des IA concrètes que l'on développe.

Comment s'est historiquement développé ce champ de l'intelligence artificielle du mouvement et du geste ?

Je dirais progressivement. On est à l'École des Mines, donc il y a deux choses. On fait de la recherche qui est orientée industrie. Cela ne veut pas forcément dire recherche appliquée, ça peut être aussi de la recherche fondamentale. Mais il faut que l'industrie soit intéressée par nos recherches. Et par ailleurs, on fait aussi de l'enseignement par la recherche. Qui dit industrie dit très souvent opérateur. Alors ça peut être un opérateur dans une chaîne de montage d'une marque automobile, mais ça peut être aussi l'opérateur d'une marque de luxe. Les indus-

tries manufacturières et les industries créatives sont intéressées, ont des défis communs par rapport à l'ergonomie au travail et à la transmission des savoir-faire. Mais le geste n'est pas le même. Dans l'automobile, on a plutôt des gestes répétitifs, avec une cadence ajustée au millimètre. Alors que dans les industries créatives c'est différent, il y a un petit peu plus d'expressivité. Mais ce ne sont pas des artisans au sens abstrait du terme. Il y a aussi des chaînes de production. Mais on trouve des différences en termes de gestes. Les gestes des mains et des doigts sont partout dans l'industrie et la convergence c'est le travail manuel. Donc j'ai commencé par les mains et les doigts. Je suis passé par le conduit vocal, ensuite je suis passé sur des gestes beaucoup plus amples. Je pense que c'est important de préciser que j'ai beaucoup plus travaillé sur les aspects cinématiques et beaucoup moins sur les aspects cinétiques, donc plutôt sur les trajectoires, les rotations, les accélérations.

Je ne suis pas dans l'IA générique. L'IA peut être appliquée sur des données issues de la finance, des assurances, des banques et du médical. Nous, ce que l'on fait au Centre de Robotique, c'est de l'IA centrée sur l'humain. Si tu tapes dans un moteur de recherche « *IA centrée sur l'Humain* » dans la plupart des cas tu vas tomber par exemple sur l'Institut de Stanford, le HAI [Institute for Human-Centered Artificial Intelligence[4]], où on traite plutôt des aspects éthiques et juridiques. Nous, ici, on donne une définition provenant de l'ingénierie, donc pour nous, c'est l'ensemble des méthodes et des pratiques d'apprentissage machine appliquées aux signaux provenant du corps humain. De plus, on a deux particularités et contraintes, mais qui font aussi le charme de notre travail par rapport à l'IA générique. Notamment, le fait que l'on travaille avec l'humain, que l'on a un accès très difficile à l'humain. C'est beaucoup plus difficile d'enregistrer une main que de télécharger des jeux de données issues de la température, de l'humidité ou des délais de trafic routier ou quoi que ce soit d'autre que de l'humain. Donc ça c'est une première contrainte forte avec toutes les questions liées au RGPD [Règlement Général sur la Protection des Données[5]]. Et la deuxième particularité, c'est le temps réel. C'est-à-dire que très souvent, on doit prendre une décision sur la façon dont l'humain est en train d'effectuer le geste en temps réel, c'est-à-dire avant même que le geste soit complété. C'est ce qu'on appelle dans notre langage, prendre une décision sur la base des données manquantes, parce que l'ordinateur n'aura pas encore tout vu avant de se décider sur le geste.

4 Le HAI est un institut de l'Université de Stanford ouvert en 2019 et dédié à l'étude et au développement de technologies et d'applications d'intelligence artificielle centrées sur l'humain [URL : https://hai.stanford.edu].

5 Le RGPD est un règlement européen de 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données [URL : https://www.cnil.fr/fr/reglement-europeen-protection-donnees].

Tu as dit que tu avais commencé à travailler sur cette question en 2005. À ce moment-là, c'était déjà un champ qui existait ? Vous étiez combien en Europe à travailler là-dessus par exemple ?

À l'époque où je faisais la bibliographie de ma thèse, il y avait très peu d'articles sur les gestes, et aucun sur les gestes de la musique. Il y avait des recherches sur les gestes des doigts parce qu'il s'agissait de systèmes qui avaient comme objectif de remplacer le clavier. Ils mettaient une caméra pour capturer le mouvement des doigts et faisaient une sorte de traduction en touche de clavier. Mais cela concernait très peu de gens, très franchement. Il n'y avait certainement pas ce qu'on a aujourd'hui, des projets et des approches plutôt complètes avec les anthropologues, les sociologues, les ingénieurs et les développeurs. À l'époque, les recherches existantes développaient des approches très uni-dimensionnelles. Les ethnologues travaillaient dans leur coin. Les ingénieurs dans leur coin. Il n'y avait pas de regards croisés. C'est au début des années 2010 qu'on a commencé tout timidement. Et bien évidemment, comme avec chaque nouveau projet innovant, les résultats sont un peu moins intéressants que ce qu'on espère au début. Aujourd'hui, il y a beaucoup de produits-étagères qui te proposent une capture. Cela peut être satisfaisant, comme cela peut ne pas être satisfaisant. Aujourd'hui, je ne fais plus ce que je faisais en 2005. Je vais beaucoup plus sur la détection, sur la modélisation des descripteurs mathématiques du geste, alors qu'à l'époque c'était plutôt : « où se trouve le doigt dans l'image ? »

Quelle forme prend la notion de « descripteur mathématique du geste » ?

Ce qui est intéressant à évoquer, c'est la méthodologie un peu générique que l'on a développée. Souvent, je la présente sous forme de recettes, une méthodologie à six étapes.

La première étape, c'est la capture du mouvement. Et qui dit captures dit capteurs. On a trois types de capteurs. Il y a les capteurs fondés sur la vision, sur les caméras. Les premiers capteurs de ce type étaient hyper chers parce qu'ils étaient employés dans l'industrie du cinéma et le biomédical. Ils coûtaient environ 50 à 60 000 € et donnaient un rendu fin avec une précision assez importante. Mais tout le monde n'avait pas accès à ce type de capteurs car il y avait une contrainte très importante. Ils étaient préinstallés dans des studios, donc on ne pouvait pas prendre le capteur et aller visiter un artisan dans un atelier. Il y a ensuite les capteurs très bas coûts, des caméras. Ils sont venus beaucoup plus tard. Là, on filme et on applique ce que l'on appelle des méthodes de traitement d'images, pour segmenter les différentes composantes de l'image et arriver jusqu'au corps humain, ses contours, ses différentes parties, les articulations, les segments. Il y a déjà une première couche d'intelli-

gence artificielle ici, qui a comme objectif d'estimer les poses humaines et d'extraire ce qu'on appelle le squelette humain. Un squelette, ce n'est pas un vrai squelette mais une sorte de pseudo squelette superposé sur la figure humaine, sur l'image. C'est un capteur avec beaucoup moins de précision mais qui nous a permis de développer plein d'applications intéressantes pour le grand public. Il a généré une dynamique, une certaine concurrence entre les chercheurs parce qu'on avait accès à des capteurs bas coûts et on pouvait faire des choses. Il y a enfin les capteurs que l'on appelle « *wearable* » et que l'on porte sur notre corps. Ces capteurs nous donnent des rotations, des accélérations. Ils sont relativement chers, entre 10 et 20 000 €. Et là tu peux capturer la tête, les gestes des doigts et ça a beaucoup d'avantages parce que ça se met dans une valise. C'est transportable. Mais tu ne peux pas avoir d'information sur la scène, alors qu'avec l'image, tu as les outils de l'artisan, tu as sa matière. Là, tu n'as que le corps humain. Ça c'est la première étape, l'enregistrement des données.

La deuxième étape est l'analyse du signal. On a une image. Comment segmente-t-on l'image ? Comment on différencie l'arrière-plan de l'avant-plan ?

La deuxième étape, c'est l'extraction des descripteurs de mouvement. C'est une sorte de variable mathématique qui est censée décrire le mouvement humain. Sur les capteurs, ces descripteurs sont déjà fournis. Tu mets le vêtement et tu as tes rotations et tes accélérations. Mais quand tu as affaire à de l'image, l'image ne fournit pas de descripteurs, il faut que tu ailles les chercher dans l'image, les concevoir, les mesurer. Donc ces descripteurs, il faut les définir en fonction de l'application. Il n'y a pas de règles là-dessus. C'est l'expérience qui nous montre que quand on a un geste plus artistique, le performeur est plutôt dans la gestion continue des accélérations et des décélérations, et dans la gestion de l'espace. Alors que quand on a affaire à un artisan, on a la manipulation des objets et on a parfois des gestes répétitifs et précis. Là, souvent, on fait appel à des descripteurs de trajectoire et de positions.

Après on a la représentation. C'est ce qu'on a proposé pour le projet européen Mingei. Dans la représentation, on a ce qu'on appelle les modèles mathématiques, la partie IA. Par exemple, on a fait une proposition pour ce projet et c'est là où il m'a semblé que ce n'était peut-être pas le bon cadre pour faire cette proposition. On a fait beaucoup d'observations et on a proposé une sorte de modélisation mathématique de la dextérité. Ça peut paraître un peu farfelu, mais on a voulu mettre les bases. Cette représentation de la dextérité avec des notions mathématiques se fait, selon nos travaux, sur la base de quatre hypothèses. 1) L'hypothèse de la transition : le geste passe par plusieurs phases, évolue dans le temps et chaque phase est corrélée à la phase précédente. 2) L'hypothèse de la synergie des différentes parties du corps : tu prends

le cas d'un artisan, il manipule sa matière et ses deux mains travaillent de manière synergétique. 3) L'hypothèse de la médiation des différentes parties du corps : afin de bouger la main, il y a une médiation sérielle entre le coude et la main, le coude influence le mouvement de la main. Mais il y a aussi une médiation non-sérielle entre l'épaule et la main comme on l'a vu dans le soufflage de verre. 4) L'hypothèse de la corrélation entre les trois dimensions de l'espace : quand on déplace sa main dans l'espace, mathématiquement, ce déplacement se traduit en trois mouvements différents, un sur chaque axe. Cette hypothèse de l'intercorrélation dit que pour que l'humain se déplace, il faut qu'il y ait trois mouvements différents qui soient corrélés.

Ces quatre hypothèses, vous les pensez en interaction ?

Bien sûr, on assemble tout sur une série d'équations. Pourquoi ? Parce qu'une équation à deux parties. La partie gauche, c'est ce que l'on cherche et la partie droite, c'est la formule mathématique. On a une équation par articulation du geste. Comme ça, toutes les articulations sont liées entre elles. On se dit : « *voilà, je veux étudier la main droite* ». Si je veux impliquer les quatre hypothèses, il y aura une première portion mathématique qui va corréler le mouvement de la main droite, qui est mon objet d'étude, avec le moment de la main gauche et avec un coefficient + l'influence du coude droit sur le mouvement de la main + l'influence de l'épaule droite sur le mouvement de la main + l'influence de la position de la main, dans les phases précédentes, sur la phase actuelle. Après, on entraîne ce modèle sur les données qu'on enregistre avec des gestes experts avec l'objectif de mesurer le coefficient de l'influence de chaque partie de notre modèle sur le mouvement de la main. Qu'est-ce qu'on peut observer de cette façon ? On a vu sur le geste du souffleur de verre que quand il tourne sa canne, il tourne la canne avec sa main gauche et qu'avec la main droite il donne forme au verre. Donc, bien évidemment l'hypothèse de la synergie est très importante. Mais on a observé que les épaules restent constantes ou bien qu'elles font un tout petit mouvement. Il y a des choses qui peuvent être observées à l'œil nu bien sûr. Je ne dis pas qu'on a extrait de nouvelles connaissances. Mais il y a des choses qui sont observables à l'œil nu et il y a des choses qui ne sont pas observables. Ou, en tout cas, on les observe mais on a du mal à les interpréter. Là, on a quand même une sorte « d'IRM du geste[6] ». On observe que, finalement, l'épaule joue le rôle de pivot pour le reste du bras afin que la main puisse prendre la pose souhaitable pour donner la forme à l'objet. Sur ce point, j'ai un peu challengé une étudiante qui soutient sa thèse au mois de mars. Je lui ai dit : « *attends, est-ce que tu as fait l'exercice inverse ?*

6 L'imagerie par résonance magnétique (IRM) est une technique d'imagerie médicale permettant d'obtenir des vues en deux ou en trois dimensions de l'intérieur du corps de façon non invasive.

Il y a une synergie entre la main gauche et la main droite, dans le sens où la main gauche influence le mouvement de la main droite. Mais est-ce que l'inverse est aussi valable ? » Du coup, non !

Mais pour en revenir à nos étapes, la cinquième étape, c'est la reconnaissance en temps réel. C'est ce qu'on a fait au Musée des arts et métiers. Il faut que l'ordinateur puisse reconnaître le geste pour déclencher le son.

Enfin, la sixième étape, c'est ce qu'on appelle la collaboration avec la machine. La machine peut être un robot. Le robot suit le mouvement de l'humain et lui donne une pièce, l'assiste. Ça peut être aussi un ordinateur. Dans le cadre du musée c'était un ordinateur et là, l'ordinateur joue le rôle du médiateur. L'objectif, ce n'est pas d'interagir avec l'ordinateur, c'est d'apprendre ou de jouer, et l'ordinateur observe l'humain et fournit les bons feedbacks, au bon moment, pour rendre l'expérience plus riche.

Tous ces mouvements deviennent finalement des équations, n'est-ce pas ? Et c'est toi qui écris les équations, ou alors tu as des formules ?

Tout ça, c'est théorique. En fait, c'est un modèle biomécanique sur la base des observations humaines que les capteurs te donnent. Ce sont des équations théoriques. Les coefficients, au début, ce sont des alpha, bêta etc., il n'y a pas de chiffres. On mesure les coefficients sur la base des données enregistrées avec des capteurs. Donc sans données on n'a rien, juste un modèle théorique.

Qu'est-ce que t'apporte la comparaison entre le monde de l'artisanat et des savoir-faire et celui de l'industrie ?

La différence c'est la dextérité. Dans l'artisanat, on a de la dextérité. Souvent dans l'industrie, on est plutôt dans l'enchaînement de tâches. On est dans le séquentiel, dans des cadences plutôt strictes. Dans l'artisanat, c'est différent. Je suis très influencé par la poterie, qui est un cas très, très riche, parce qu'on touche directement la matière. C'est l'application la plus difficile que j'ai jamais étudiée. Et je pense que ça mérite encore une dizaine de thèses avant de sortir des conclusions qui tiennent à peu près la route. Parce qu'on a affaire à un autre objet qui tourne, qui est déformable, avec les doigts qui rentrent dans la matière.

Dirais-tu que cette place centrale de la dextérité dans les pratiques artisanales aide à penser différemment la sérialité de l'industrie ? Ou est-ce que c'est aussi dans les deux sens ? Est-ce que la sérialité de l'industrie peut aussi apporter quelque chose ?

S. M. : Déjà, la différence, c'est que le modèle que je t'ai décrit, il peut être appliqué à toutes sortes de gestes. C'est juste que dans l'industrie, il y a beaucoup

moins de choses cachées. Les choses cachées, c'est au niveau de l'ergonomie. L'industrie s'intéresse beaucoup à l'ergonomie au travail, parce que des opérateurs se font mal. Je peux te citer un exemple d'un opérateur de l'industrie automobile qui, à cause d'un problème d'épaule, se met en arrêt maladie : ça peut coûter jusqu'à un million d'euros par mois. Et ce n'est pas la sécurité sociale qui coûte cher, c'est le coût du recrutement et des formations. En fait, ce sont les besoins qui sont différents. L'industrie a besoin de définir ce qu'est le bon geste, ce qu'est le mauvais geste et de pouvoir les différencier. Parce que l'opérateur, il s'auto-challenge afin de découvrir de nouveaux chemins qui lui permettent d'aller encore plus loin dans sa performance, dans la qualité du produit qu'il fait et il se permet très souvent de prendre des risques pour son corps. Donc les risques ergonomiques et le challenge ergonomique sont bien plus forts dans l'industrie. Dans l'artisanat, c'est différent, la dextérité est bien plus riche. La créativité est plus importante. Ils peuvent aller chercher des résultats et des créations bien plus différentes que ce qu'ils avaient en tête. Dans l'industrie, on attend que ce sac qui est apprécié soit produit. La créativité est beaucoup moins sur le produit que sur la façon dont l'opérateur arrive à effectuer une certaine tâche.

Tu disais au début que peut-être, l'apport de tout ce champ serait de libérer, dans le monde industriel, les opérateurs de tâches répétitives. Qu'ils aient plus un travail basé sur la créativité et fassent des choses moins répétitives. Tes recherches sur l'artisanat te servent-elles à comprendre les relations entre séquençage et répétition ? Les gestes manuels t'aident-ils à saisir le moment où une action bascule dans le répétitif ou est une adaptation à une situation inédite ?

Cette composante répétitive est un peu prise en considération dans notre modèle théorique, à travers l'hypothèse sur la transition des différentes phases. C'est quoi la répétitivité ? C'est le fait de répéter encore et encore le même mouvement primitif. Donc on n'est plus dans la transition, car la transition peut avoir deux branches en transit vers la même phase, donc soit on répète la même phase, soit on transite vers la phase suivante et donc la répétitivité est prise en considération dans notre modèle. Donc on peut l'observer.

Est-ce que voir un artisan te fait te poser des nouvelles questions pour l'industrie et inversement ? Est-ce que bien connaître par exemple cette question de la sérialité du geste industriel permet de poser d'autres questions vis-à-vis de l'artisanat ?

Moi, j'ai travaillé sur trois niveaux. Il y a l'industrie, donc les opérateurs, il y a les artisans mais aussi les performeurs et les musiciens ; c'est encore un autre niveau. Alors jusqu'à présent, tous les gens de la communauté des chercheurs

se disent : « *ce serait bien qu'on puisse comparer !* » On n'a jamais concrètement comparé. En tout cas, on a comparé de manière qualitative, mais à ma connaissance jamais de manière quantitative. Je pense que ce serait bien de le faire. Après, les différences ? Dans l'art, la musique et la danse, ce n'est pas pareil. Il y a des points communs. Tu peux dire que l'équivalent de la partition musicale, c'est la chorégraphie. Mais le geste musical a comme objectif de déclencher un mécanisme intermédiaire qui va par la suite créer un son. Toucher la touche du piano, activer le petit marteau qui va taper sur la corde… Alors qu'avec la danse c'est moins sémantique. Je reviendrai un peu sur ce que l'on a évoqué tout à l'heure : la gestion des accélérations et des décélérations. Elle est bien plus importante, même la fréquence, dans la performance artistique que dans l'artisanat et beaucoup moins fréquente dans l'industrie. Dans l'industrie tu as beaucoup de répétitivité. Dans les métiers d'art beaucoup moins et dans l'art du spectacle très, très peu. Pareil pour la précision, parce qu'il y a aussi la notion de précision. Un opérateur doit être très précis parce que quand l'opérateur n'est pas précis ça pose problème. Un artisan peut se permettre de ne pas être précis. Dans l'art musical, ne pas être précis, je dirais, ça peut ne pas poser de problème, ça permet de chercher de nouvelles formes de création musicale. Tu prends l'exemple de la musique tonale avec la musique atonale. Dans la musique atonale, j'ose dire, c'est beaucoup moins précis et ça ne pose pas de problème.

On en vient donc forcément aux usages de ces gestes et aux contextes dans lesquels on les effectue ?

Il y a aussi la notion des outils, de la manipulation des outils. C'est très important. Les outils du potier par exemple. Je mets de côté le four et le tour. Il y a un tout petit outil qui lui permet de donner forme à son objet alors que dans la menuiserie, c'est différent. Pour chaque métier d'art, il y a plein de déclinaisons et il n'y a pas de règle, tout est permis je pense. Il y a un élément aussi très important, c'est le contexte de la transmission et si elle se fait en personne. Parce que les artisans ont beaucoup plus souvent l'habitude de travailler entre maître et apprentis que dans l'industrie. Et est-ce que l'artisan est capable de verbaliser ce qu'il fait ?

Dans le nouveau projet, en plus du geste, mon ambition est de mettre aussi la parole. De demander à la personne, qu'elle soit experte ou en apprentissage, de verbaliser ce qu'il est en train de faire. Jusqu'à présent, je n'avais pas la parole, mais là, je suis convaincu que la parole il faut qu'elle soit une des modalités enregistrées. Donc encore un capteur si tu veux ! Je pense que les artisans qui ont l'habitude de travailler avec des apprentis se sont confrontés à des challenges différents par rapport à des artisans qui travaillent seuls dans leur atelier, mais aussi par rapport à des opérateurs. En 2015, je suis allé visiter des ateliers d'une maroquinerie de luxe. Avant cela, j'étais plutôt habitué aux opérateurs de l'auto-

mobile. Dans l'automobile, chaque opérateur a son poste. J'ai appris que dans la chaîne de production de cette maroquinerie, tous les opérateurs étaient censés être capables de prendre n'importe quel poste. Donc ils étaient formés sur tous les postes de travail. Ça change quand même pas mal de choses. Et le formateur, à chaque fois, c'était le meilleur des opérateurs. Donc, dans cette maroquinerie, la notion de formateur et d'opérateur était brouillée car le formateur était aussi opérateur. Donc plus la personne est capable de verbaliser et de décrire le geste, plus riche est l'information et la connaissance pour nous les chercheurs. Je ne dirais pas que c'est mieux pour lui, pour le métier, mais en tout cas pour nous, les chercheurs, c'est beaucoup plus riche et on a un accès beaucoup plus direct avec le savoir-faire. Je ne vais rien t'apprendre là-dessus !

Quels sont les futurs enjeux dans ton champ ?

Pour que les choses soient bien claires, ce qu'on a déjà fait, ce n'est qu'une petite poussière devant la montagne de la dextérité humaine. Le geste expert est tellement riche que parfois ce que l'on fait peut paraître ridicule, mais on essaie de mettre quand même certaines bases. Jusqu'à présent, j'avais une approche où mon hypothèse scientifique partait d'une interrogation du genre : est-ce que je pourrais appliquer cette technologie ? Là, j'ai plutôt mes réponses : oui, je peux l'appliquer. Oui, ça peut donner des bons résultats. J'ai plus d'expérience, je sais à peu près où ça peut marcher, où ça ne peut pas marcher. Mais là, l'objectif est de passer à l'étape suivante. C'est-à-dire de travailler beaucoup plus sur une sorte de hiérarchie du geste, alors que jusqu'à présent j'ai travaillé sur un vocabulaire plutôt petit. Et là on peut commencer à avoir des regards croisés, dans le sens où cette hiérarchie, telle que je la conçois, peut être représentée sur une forme de pyramide, où, à la base, il y a ce qu'on appelle le « mouvement primitif ». Des micro-portions de mouvements qui peuvent être horizontaux par rapport à des métiers différents ou par rapport à des catégories de métiers. Par exemple visser, ce n'est pas exclusif à tel ou tel métier. Attraper une pièce, donner un coup de marteau. En revanche, donner un coup de marteau dans l'automobile ou donner un coup de marteau dans la menuiserie, ce n'est pas pareil. J'aimerais créer une sorte de collection de mouvements primitifs puis après considérer des gestes ayant une sémantique de plus, puis, à un niveau encore plus élevé, au sommet de la pyramide, il y aurait l'activité qui pourrait être liée au métier. Parce que jusqu'à présent, il y avait une seule couche.

Dans notre intelligence artificielle il n'y avait pas de hiérarchie. Je pense que ce concept de hiérarchie peut ouvrir le champ pour étudier les gestes de façon non fragmentée. Ce qu'on a fait dans Mingei, ce qu'on a fait par le passé, c'étaient des études fragmentées. Là, je

ne veux plus faire d'études fragmentées. Bien sûr on va étudier chaque métier différemment, mais on va commencer à comparer et à envisager la création d'une IA qui pourrait être utile. Parce que l'IA n'est pas que le modèle mathématique, il y a aussi l'algorithmique. Le modèle mathématique que je t'ai décrit peut être appliqué à beaucoup de choses, mais ce n'est pas le même algorithme. L'algorithme est différent. Les données sont différentes. Les données seront toujours différentes, mais l'algorithme ? Est-ce qu'il y a des utilisations d'un même algorithme entre les différents métiers ?

En termes de recherche, un autre challenge est de créer un conservatoire de gestes. On a posé les bases avec le projet Mingei. On est passé par les archives numériques dans le sens texte et image pour y ajouter aussi toutes les données enregistrées avec les capteurs. Mais il faut maintenant commencer à avoir une quantité considérable parce que là, avec les quelques enregistrements à droite à gauche, ça reste relativement pauvre. Il me semble que c'est quelque chose d'important et qui va aussi permettre à de nouvelles études de voir le jour. Je le vois comme une sorte de contribution vers d'autres disciplines de recherche. Je ne sais pas, mais toi, si tu avais à ta disposition des données facilement accessibles avec le petit avatar qui bouge en plus de tes images, peut-être qu'en termes d'ethnographie ça pourrait ouvrir d'autres champs ? Alors que l'algorithme t'intéresse beaucoup moins j'imagine ?

Cette hiérarchisation dont tu parles, d'aller dans la finesse ou les typologies de gestes, est-ce qu'il va falloir aussi développer des nouveaux moyens de capter ou des nouveaux modèles mathématiques ou s'agit-il plutôt d'une façon de recombiner des choses qui existent déjà ?

Les capteurs sont les mêmes, c'est la même capture. Le problème aujourd'hui, dans le milieu de la science des ordinateurs, ce ne sont pas les capteurs, les composantes électroniques. Ce qui manque, c'est la partie algorithmique. On ne sait pas trop quoi faire avec le capteur qu'on a. C'est pour ça que le labo n'est pas positionné sur la robotique mais sur ce que l'on peut faire avec la robotique. Donc ce sont les algorithmes qui manquent.

C'est pour ça qu'il faut multiplier les cas ?

Oui, parce que je parle de la poterie, mais ce qu'on a étudié, c'était quoi ? C'est la création d'un saladier, d'un bol ! Ça peut paraître ridicule. Donc il nous faut beaucoup plus d'expérience quand même. Je fais partie des gens qui, du côté de l'ingénierie, ont eu quand même la chance d'avoir plusieurs expériences. Je ne me plains pas ! C'est juste que maintenant que j'ai à peu près compris et que j'ai trouvé mes repères, je pense qu'il faut quand même passer à l'étape suivante et croiser des hypothèses, croiser des concepts, concevoir des IA plus

aptes et mieux équipées pour qu'elles soient mises au service de l'humain. Parce que pour moi, l'IA ce n'est pas pour remplacer l'humain. Je le souligne quand même.