



UNIVERSITÉ
DE LORRAINE

UFR MATHÉMATIQUES
ET INFORMATIQUE



MASTER 1 SCA 2012-2013

PROJET TUTORÉ

Nancy au doigt et à l'oeil grâce au dispositif D-MIN

Auteurs :

Amaury L'HUILLIER
Florian MARCHAL

Encadrant :

M. Maxime AMBLARD

3 juin 2013

Master 1 SCA
UFR Mathématiques et Informatique
Rapport de projet tutoré

Tuteur : Amblard Maxime
Encadrant projet D-MIN : Castagnos Sylvain
Organisme d'accueil : LORIA, équipe KIWI

Nancy au doigt et à l'oeil grâce au dispositif D-MIN

L'Huillier Amaury et Marchal Florian

Nancy, le 3 juin 2013

Table des matières

Introduction	3
1 Interaction et moyen d'interaction	4
1.1 L'interaction	6
1.2 Les moyens d'interaction	9
1.2.1 Les plus anciens : Clavier, souris, joystick et gant de données	9
1.2.2 Les plus récents : Kinect, Eyetracker, Surfaces tactiles et commande vocale	12
1.2.3 Les Autres	17
1.2.4 Conclusion	19
2 L'interaction gestuelle	20
2.1 Qu'est ce qu'un geste?	20
2.2 Les gestes et la Kinect	21
3 La Kinect et Street View	23
3.1 L'application	23
3.1.1 Google Street	23
3.1.2 Les choix techniques	25
3.1.3 Description des classes	26
3.2 La collecte des gestes	27
3.3 Une course folle	30
3.4 Résultats et analyse	32
Bilan personnel	35
Annexes	37
.1 Analyse chi2	37
.2 Relevé de gestes et questionnaire	41

Introduction

Dans un monde où la technologie et les possibilités offertes par cette dernière sont en plein essor, de nouveaux dispositifs d'interactions émergent et tendent à remplacer le clavier et la souris qui pendant longtemps ont été les seuls dispositifs d'interactions avec les applications numériques. Le Track pad ainsi que les surfaces tactiles sont les précurseurs de ce changement. Cependant une des plus grandes innovations dans l'art de l'interaction fut la sortie de la Kinect en 2010, qui remporta un succès unanime au point de figurer dans le Guinness Book pour le titre de « l'accessoire high-tech le plus vendu dans un court laps de temps ». Conçue initialement pour des jeux vidéo, elle devient un moyen d'interaction à part entière pour de nombreuses applications.

La Kinect, n'est que l'un de ces nouveaux dispositifs d'interaction. En effet, nous ne pouvons nous permettre d'aborder le sujet des nouveaux moyens d'interactions sans citer des dispositifs tels que le contrôle vocal (qui est également présent dans la Kinect), les surfaces tactiles, les dispositifs oculométriques (plus couramment appelé eyetrackers) ou encore les nouveaux dispositifs très prometteur actuellement en cours de développement, à l'image du Leap Motion ou encore du bracelet MYO, dispositifs dont nous reparlerons dans la première partie de ce rapport.

De nouveaux moyens d'interactions sont créés au fil des progrès technologiques, cependant peuvent-ils tous se substituer au clavier et à la souris si présents et ancrés dans les mœurs ?

Sont-ils tous aussi efficaces par rapport à une tâche donnée ?

Chaque moyen d'interaction doit-il être utilisé dans un cadre précis, ou plutôt, chaque tâche ne possède-t-elle pas un moyen d'interaction qui lui est le plus adapté.

Certains moyens d'interactions peuvent-ils être inefficaces et désagréable à utiliser pour une tâche donnée ?

En tant qu'étudiant en sciences cognitives, nous avons décidé de nous intéresser aux nouveaux moyens d'interactions car nous pensons qu'à l'heure actuelle, il est important d'aborder le sujet des nouveaux moyen d'interaction avec un certain recul, c'est-à-dire en laissant de côté les aspects ludiques et attirants pour se focaliser sur une problématique centrée sur l'utilisateur. Ce rapport n'a pas pour but de définir les meilleurs gestes pour une interaction ou le meilleur moyen d'interagir avec les nouveaux moyens d'interactions mais de réfléchir sur cette problématique en se focalisant sur l'utilisation d'une Kinect pour naviguer dans Street View.

Notre projet tutoré s'inscrit dans la cadre de renaissance 2013 Nancy, un événement durant lequel nous avons eu la chance de promouvoir notre filière et de faire découvrir quelques domaines d'applications des sciences cogni-

tives. Pour cet événement, nous avons réalisé une application permettant à un utilisateur de se déplacer dans Street View à l'aide d'une Kinect. Nous avons décidé de centrer notre problématique sur la question suivante : La Kinect est elle un moyen d'interaction adapté pour se déplacer dans Street View ?

Avant d'aborder en profondeur cette question, une première partie de ce rapport sera consacrée aux nouveaux moyens d'interaction et aux principes de l'interaction. L'interaction étant au cœur de notre projet il convient de définir clairement ce concept et de faire un historique de ce qu'a été l'interaction au fil des années et de suivre son évolution par rapport à la technologie. Nous expliciterons brièvement le fonctionnement et les paradigmes d'utilisation de dispositif tels que le joystick, l'eyetracker, le Leap Motion, les surfaces tactiles, la reconnaissance vocale et bien sur, la Kinect qui sont considérés comme étant des nouveaux moyens d'interaction.

Utilisant la Kinect et sa fonction de reconnaissance gestuelle dans notre application, nous consacrerons la deuxième partie de ce rapport à la notion de geste et à son utilisation pour interagir avec une Kinect.

L'application que nous avons créée pour permettre à un utilisateur de se déplacer dans Google Street sera présentée dans la dernière partie de notre rapport. Nous y décrirons le cahier des charges et nous détaillerons le processus que nous avons suivi pour arriver à obtenir ce qui nous était demandé. Nous expliquerons donc les aspects techniques, c'est-à-dire le choix des langages de programmation que nous avons utilisés et la prise en main du kit de développement pour la Kinect. Nous exposerons aussi les problèmes que nous avons rencontrés ainsi que la manière dont nous les avons résolus. Cette dernière partie reviendra sur le protocole de test que nous avons mis en place, afin de tenter de répondre à notre problématique. Nous expliciterons dans un premier temps le protocole déployé afin de collecter les gestes sur lesquelles nous appuierons notre étude. Dans un second temps, nous présenterons résultats obtenus afin de les comparer à nos hypothèses formulées précédemment.

1 Interaction et moyen d'interaction

Comme nous l'avons mentionné dans l'introduction, les technologies de l'interaction ne cessent d'évoluer et permettent de créer de nouveaux cadres d'interactions. Ces évolutions et progrès sont au cœur d'une discipline née dans les années 1950, les sciences cognitives. Cette discipline rassemble plusieurs disciplines qui sont la philosophie, la linguistique, l'anthropologie, les neurosciences, l'informatique et la psychologie. Ces six disciplines ne furent

pas toutes présentes en tant que telles lors de la naissance des sciences cognitives. En effet, c'est le rapprochement de discipline telles que l'intelligence artificielle, la psychologie et la biologie qui ont composé l'essence des sciences cognitives. Les sciences cognitives étant la résultante des liens interdisciplinaires existant entre plusieurs disciplines majeures G.A.Miller, un des pères de la psychologie cognitive, propose de représenter les liens existants entre ces six composantes majeures à l'aide d'un hexagone.

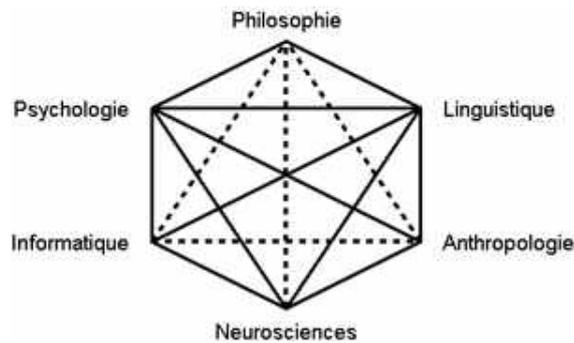


FIG. 1 – Diagramme des sciences cognitives

Sur la figure 1 les traits pleins représentent les liens établis lors de la création des sciences cognitives alors que les pointillés représentent les liens qui n'ont pas encore pu être établis. Le terme de lien signifie qu'il y a un rapport direct, des interactions, des paradigmes communs entre deux disciplines. Par exemple, le lien entre neurosciences et informatique est en trait plein car les deux disciplines possèdent une discipline commune : l'intelligence artificielle. Les sciences cognitives sont très présentes dans le monde actuel où la technologie est en pleine expansion. Elles apportent une vision centrée sur l'homme et donc dans le cas des nouvelles technologies centrées sur l'utilisateur. Dans le cadre de ce projet tutoré, notre travail se situe dans la branche des sciences cognitives que l'on appelle les IHM, acronyme pour Interface Homme Machine, à ne pas confondre avec Interaction Homme Machine¹. Bien que ces deux mots soient intimement liés, notre travail se situe plus au niveau de l'interaction plutôt que de l'interface en elle-même. Nous allons donc dans cette première partie nous intéresser au concept d'interaction puis nous ferons un aperçu des dispositifs d'interactions existant tels que ceux cités dans l'introduction.

1. L'Interaction Homme machine est la science de l'interaction alors que l'Interface Homme Machine est l'ensemble de dispositifs matériels et logiciels permettant à un utilisateur de communiquer avec un système informatique

1.1 L'interaction

La définition du terme interaction telle qu'elle peut apparaître dans un dictionnaire est la suivante : « Réaction réciproque de deux phénomènes l'un sur l'autre »². Ces quelques mots nous en disent déjà beaucoup sur ce qu'est une interaction, cependant, nous ne pourrions nous en contenter pour ce qui est de notre travail. A l'exception de toutes les autres interactions existantes, les interactions homme-machine sont les seuls cas d'interactions où la communication entre les deux entités doit se réaliser au moyen d'une interface utilisateur. Celle-ci permet à l'homme et à la machine de communiquer dans un langage qu'ils partagent. Les interfaces utilisateurs sont donc la charnière pour les interactions homme machine.

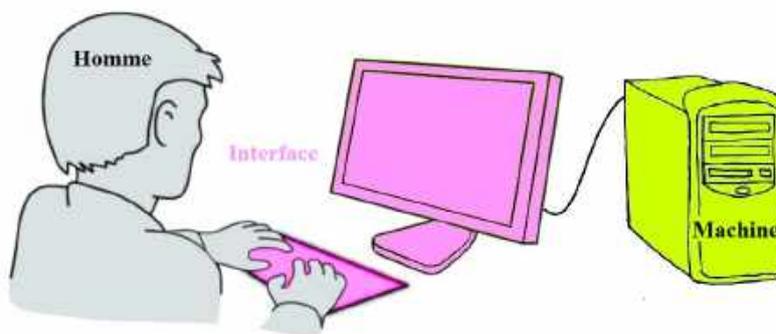


FIG. 2 – L'interface homme machine

Une interface homme machine est par définition³ « l'ensemble des dispositifs matériels et logiciels permettant à un utilisateur d'interagir avec un système interactif ». L'IHM a pour rôle de faire communiquer un homme et une machine malgré le fait qu'ils ne parlent pas le même langage. Au temps des premières machines informatiques, la communication s'effectuait via des transistors et de nombreux réseaux de câbles. Il n'y avait pas d'interface graphique ni d'écran ce qui exigeait donc une certaine maîtrise technique pour communiquer avec une machine. Par la suite, la communication se faisait via des cartes perforées, puis via des lignes de code, ce qui a amené à l'utilisation du clavier informatique. Le domaine des interfaces homme machine fut renversé par Douglas Engilbert, qui proposa les interfaces graphiques de type WIMP, acronyme pour « Windows, Icône, Menu, Pointeur ». Il s'agit du type d'interface graphique que nous connaissons tous et qui est présent sur tous les systèmes d'exploitation, c'est dire une fenêtre dans laquelle se trouve une application, un pointeur que l'on déplace avec une souris pour

2. <http://www.larousse.fr/dictionnaires/francais/interaction/43595?q=interaction#43518>

3. http://www.journaldunet.com/encyclopedie/definition/263/52/22/interface_homme_machine.shtml

interagir avec les icônes et le menu de cette fenêtre.

Dans la définition du dictionnaire citée précédemment, il est question de « phénomène », ce terme fait référence aux acteurs de l'interaction, car comme définit, une interaction ne peut avoir lieu s'il n'y a pas au moins une paire d'entité. Ainsi on ne peut pas dire qu'une entité isolé soit en interaction⁴. Il est intéressant de remarquer l'utilisation du terme « phénomènes », cela nous rappelle que l'interaction n'est pas qu'un terme anthropomorphe et donc que des interactions peuvent se produire entre des entités non vivantes. Par exemple, deux planètes qui s'attirent sous l'effet de l'attraction sont en interaction, tout comme un homme qui interagit avec un ordinateur. Le terme de « phénomène » est important, il sous-entend l'idée que l'on ne parle pas simplement d'une entité mais d'une entité qui réalise une action. De plus le terme « phénomène » permet aussi de rendre la définition valable pour les interactions physiques telles que l'exemple des planètes cité précédemment où l'interaction entre les planètes ne résulte pas d'une volonté ou d'un désir mais de la force d'attraction, de leur distance et de leur masse, d'où le terme de phénomène.

Si l'on parle d'interaction, c'est qu'il y a action de la part des entités concernées par cette interaction, c'est-à-dire que dans le cas d'une interaction homme machine, les deux acteurs agissent, tant l'homme que l'ordinateur. Dans le cadre des IHM, il existe des actions de différentes natures : lorsqu'une personne appuie sur les touches d'un clavier pour produire du texte, lorsque l'on appuie sur une surface tactile, lorsque l'on parle à un dispositif de reconnaissance vocale etc. Des actions d'une même nature peuvent prendre différentes formes, par exemple, dans le cas de la commande vocale, la nature de l'action émise par la personne est la voix, par contre les mots ou les phrases émises peuvent produire des actions différentes. Une interaction est en permanence suivit d'un effet, c'est le cas lorsqu'après avoir appuyé sur une touche d'un clavier un homme voit une action s'effectuer, la machine peut donc effectuer une action logiciel telle que sauvegarder un document, quitter une fenêtre ou afficher à l'écran le caractère correspondant à la touche enfoncée.

4. En dehors de l'interaction qu'elle entretient avec son environnement auquel cas l'environnement devient un acteur de l'interaction, ce qui ne contredit en rien ce qui vient d'être dit.

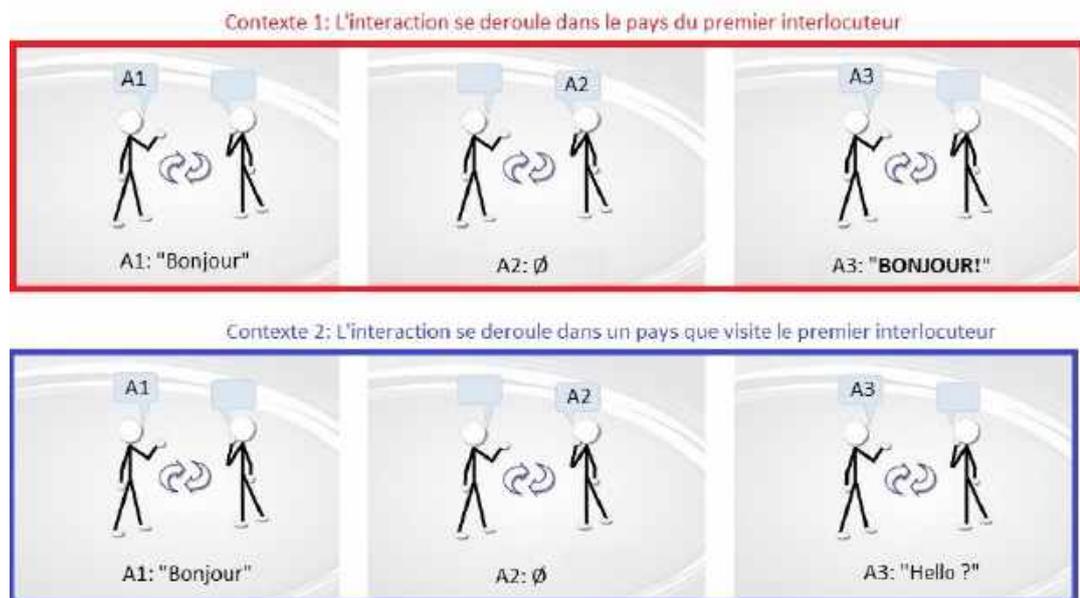


FIG. 3 – Le contexte dans l'interaction

Lors d'une interaction, une des deux entités produit une action A1, cette action a pour conséquence de modifier l'état de l'autre entité qui va produire en réponse une action A2. Cette dernière va informer la première entité que son action a été perçue. Cependant l'action A2 produite en réponse à A1 peut très bien ne produire aucun effet. C'est-à-dire que l'entité ayant reçu A1 ne produit aucune action (on considère que ne pas produire d'action peut être entendu comme une action). La figure 3, illustre justement ce cas, le premier interlocuteur dit bonjour à un second, cependant selon le contexte, le silence de ce dernier peut signifier plusieurs choses : soit qu'il ne parle pas la langue, soit qu'il n'a tout simplement pas entendu parler le premier protagoniste, d'où les réactions différentes de ce dernier. On pourrait alors se demander comment distinguer une action "nulle" (faite dans le cas d'une interaction) d'une "non" action. Pour répondre à cette question, nous devons ajouter une notion supplémentaire qui n'est pas dans la première définition, la notion de cadre. En effet, une interaction se déroule toujours dans un cadre défini. Un cadre n'est pas nécessairement matériel ou physique, le cadre d'une interaction fait référence à toutes les « connaissances » que les deux entités doivent posséder. Si l'on reprend l'exemple de la personne qui effectue une saisie clavier, le cadre de l'interaction définit le contexte dans lequel se trouve cette personne. Admettons que la touche enfoncée soit la touche « z », si la personne est en train de rédiger un texte alors l'appuie de cette touche aura pour effet d'afficher à l'écran le caractère « z », cependant si cette personne joue maintenant à un jeu vidéo, l'appuie de cette même

touche aura pour conséquence de faire avancer son personnage dans le jeu. Il y a donc autant d'actions possibles pour une touche donnée qu'il y a de cadres différents. C'est donc pour cela que le cas où l'action A2, aussi « nulle » soit elle, peut avoir une signification très importante selon le cadre de l'interaction. Dans le cas où cette action A2 est émise par un logiciel en réponse à une action A1 faite par un utilisateur, A2 peut soit signifier un problème logiciel, soit signifier que tout se déroule bien etc.

1.2 Les moyens d'interaction

La notion d'interaction étant définie, nous allons donc dans la partie suivante nous intéresser aux différents moyens d'interaction, cette partie n'a pas pour but de faire la liste de tous les moyens d'interaction existant mais plutôt d'en démontrer la diversité et les paradigmes d'utilisation qui les définissent. Nous ferons également une brève présentation des caractéristiques matérielles de chacun d'eux, expliquant ainsi les spécificités de la Kinect.

1.2.1 Les plus anciens : Clavier, souris, joypad et gant de données



FIG. 4 – figure 1 : Un clavier

Le clavier, après les systèmes de cartes perforées est le premier moyen d'interaction avec la machine informatique à être mis au point. Présenté dans les années 1960, son fonctionnement est identique à celui de ses prédécesseurs présents dans les machines à écrire et les téléscribes. C'est-à-dire qu'on y retrouve les vingt-six lettres de l'alphabet ainsi que d'autres touches servant à écrire des caractères spéciaux ou à effectuer certaines actions spécifiques selon le cadre de l'interaction, c'est le cas notamment pour les touches situées en haut du clavier. Le clavier peut posséder un arrangement des lettres différent selon les pays, sur la figure 4 il s'agit d'un clavier AZERTY, l'autre version la plus courante étant le clavier de type QWERTY.



FIG. 5 – Une souris informatique

Mise au point quelques années après le clavier, la souris informatique fait son apparition publique en 1968 (bien qu'inventée cinq années auparavant). On attribue souvent à tort, l'invention de la souris à la société Apple (qui la rendit tout de même célèbre), mais son créateur est Douglas Engelbart qui mit au point un premier prototype de souris, appelé à l'époque « trackball » dû à la présence d'une boule enfermée dans une boîte en bois. Bien que son esthétique et sa technique (capteur optique, laser, sans fil...) aient évolués au cours des années, la souris que nous connaissons à l'heure actuelle repose sur le même principe, c'est à dire qu'un déplacement de la souris sur une surface, se traduit par une action de déplacement défini par le programme dans lequel la souris est utilisée. La figure 5 représente une souris "classique" composée de deux boutons le clic gauche et le clic droit mais comme nous l'avons dit, celle ci peut offrir de multiples autres options (boutons supplémentaires, réglage de la taille et/ou de la sensibilité...). Dans son usage le plus courant, la souris a pour but de déplacer le pointeur et d'interagir avec des fenêtres via les boutons positionnés sur sa face supérieure, mais elle peut également servir à orienter une caméra ou à déplacer un personnage dans un jeu vidéo.

Le clavier et la souris sont donc les deux périphériques ayant rendu possible la création et l'utilisation des interfaces homme machine de type WIMP. Bien qu'ils soient les moyens d'interaction avec un ordinateur les plus répandus et les plus utilisés à l'heure actuelle, d'autres périphériques sont également utilisés afin de réaliser des actions où clavier et souris montreraient leurs limites.

Cependant, la présence si répandue du couple clavier-souris oriente notre définition de l'IHM et de ce fait il est difficile de sortir de ce paradigme d'interaction. C'est pourquoi d'autres dispositifs ont été mis au point même si pour la plupart, il ont du mal à dépasser le cadre "WIMP".



FIG. 6 – Un joypad

Inventés au début du XX^{ème} siècle, les premiers joysticks étaient destinées à une toute autre application que l'informatique, puisque ces derniers furent mis au point dans le domaine de l'aviation. En effet, le joystick, ancêtre du joypad est le « manche à balais » qui sert à piloter un avion, alors que le joypad, désigne le dispositif utilisé en informatique, le plus souvent pour des jeux vidéo. Tout comme une souris, le joypad est constitué de boutons permettent de réaliser des actions selon le contexte d'utilisation. La partie centrale également appelée « manche » est mobile et permet d'effectuer des déplacements. Le manche est souvent constitué d'un système de retour à sa position d'origine lorsqu'aucune force n'est exercée sur ce dernier, ce qui bien évidemment le rendrait tout à fait inutile pour déplacer un pointeur dans un système d'exploitation. Cependant, il peut s'avérer très utile pour des tâches telles que la simulation aéronautique dans son utilisation classique ou encore dans des applications de visualisations en trois dimensions, les possibilités de rotations offertes par la partie mobile étant très appréciées.



FIG. 7 – Les gants de données pirigrine

Bien qu'il rentre dans le cadre de l'interaction gestuelle, les gants de données à l'inverse de la Kinect font partie des dispositifs devant être porté par l'utilisateur. Ces derniers ne détectent que les mouvements effectués par les

mains, cependant ils offrent une meilleure précision et un temps de réponse inférieur à la Kinect. Ce gain de performance s'explique par l'électronique qui est présente quasiment partout sous la surface du gant permettant donc un traitement du signal moins bruité que dans le traitement vidéo. Actuellement, les gants de données utilisent des capteurs électroniques détectant la pression s'appliquant à l'intérieur mais aussi à l'extérieur du gant selon le mouvement effectué. Avant les capteurs, le premier gant de données mis au point par Thomas Zimmerman en 1981-1982 fonctionnait à l'aide de fibres optiques présentes dans les gants de telle sorte à ce que ces dernières soient positionnées le long des doigts. Elles permettaient ainsi, suite à leurs déformations, de rendre compte des articulations utilisées et donc du geste effectué. Les gants de données précurseurs du mouvement de la réalité virtuelle, ont par la suite évolués et permettent à l'heure actuelle d'effectuer des manipulations d'objets en trois dimensions grâce notamment à la préhension virtuelle rendu possible par la précision de ce dispositif.

Les gants de données doivent être portés par l'utilisateur ce qui en fait un dispositif intrusif pouvant gêner l'utilisateur et apporter diverses contraintes (transpiration, contrainte de déplacement si les gants sont reliés à l'ordinateur...).

1.2.2 Les plus récents : Kinect, Eyetracker, Surfaces tactiles et commande vocale



FIG. 8 – La Kinect

La Kinect, comme nous l'avons expliqué dans l'introduction, est un périphérique d'interaction développé par Microsoft, plus précisément le côté logiciel est réalisé par la société Rare, alors que le côté matériel est l'œuvre de la société PrimeSense. Le nom Kinect vient de la contraction du mot « cinématique » faisant référence à la cinématique des mouvements corporels mis en œuvre pour interagir avec la Xbox et du mot « connexion » faisant référence au côté immersif offert par la Kinect où l'utilisateur est directement connecté à son application. Bien qu'elle ne fut pas le premier dispositif d'interaction gestuel proposé au grand public,⁵ de par la technologie qu'elle embarque et les possibilités qu'elle offre, la Kinect remporte un grand succès à sa sortie.

5. l'EyeToy sortie pour la PlayStation 2 sortie en 2003 offrait déjà la possibilité d'utiliser le geste pour jouer, bien que réadaptée pour la PlayStation 3, le dispositif n'a que très peu fait parler de lui.

Au départ conçue pour des jeux vidéo sur la Xbox 360, certains voient en elle un réel potentiel. A tel point que le dispositif logiciel se voit piraté, permettant ainsi à des développeurs d'exporter la Kinect sur des ordinateurs afin d'utiliser ce matériel pour de toutes autres interactions. Apparaissent alors des vidéos amateurs où la Kinect remplace la souris, permettant de contrôler son ordinateur (lancer des applications, monter ou descendre le son d'une piste audio, mettre sur pause une vidéo) grâce à la commande gestuelle ou vocale. Suite à cela, Microsoft réagit et décide de créer un kit de développement (SDK, acronyme Software Development Kit, encore appelé devKit) gratuit permettant ainsi à tout possesseur d'une Kinect de donner libre cours à son imagination et de dépasser le cadre de l'interaction clavier souris telle que nous la connaissons à l'heure actuelle. Le kit de développement permet ainsi de prendre en main les fonctions algorithmiques nécessaires pour développer ses propres applications.

Comme il est possible de le voir sur la figure 8, la Kinect dispose de trois capteurs qui sont un laser à infrarouge, une caméra permettant de mesurer la profondeur et une caméra RGB. Posée sur un pied motorisé, cette barre horizontale connectée peut s'incliner de plus ou moins 27 degrés, lui permettant de s'adapter à différentes situations. Un des grands avantages de ce périphérique est qu'il ne nécessite pas de calibrage particulier, il s'adapte à tous les environnements et grâce à ses capteurs infrarouges il peut fonctionner dans des conditions de faible luminosité. Il faut cependant observer une contrainte de distance minimale et maximale entre la Kinect et les personnes l'utilisant. Bien que la Kinect puisse détecter quatre personnes situées dans son champ de vision, seules deux personnes peuvent l'utiliser de manière simultanée. Pour détecter une personne et donc ses gestes, la Kinect envoie des ondes infrarouges exactement de la même manière que le sonar d'un sous marin afin de détecter la présence ou non d'une personne. Des algorithmes utilisant les informations récupérées par les différents capteurs sont chargés de déterminer si oui ou non il s'agit d'une personne. La Kinect dispose également d'un microphone multi-réseau, permettant comme nous l'avons dit la reconnaissance vocale dans plusieurs langues.

Bien que polyvalente en offrant reconnaissance vocale et gestuelle, la Kinect possède des faiblesses telles qu'une précision limitée dans la reconnaissance des mouvements fins et une vitesse de réaction faible pouvant nuire aux tâches les plus exigeantes. Ces points faibles sont en partie dû au matériel mais surtout aux algorithmes tels que celui reconnaissant les doigts de la main qui ne sont pas fournis par le SDK, même si certains "amateurs" y sont parvenus avec plus ou moins de succès. Les deux vrais points faibles de la Kinect résident dans l'espace nécessaire afin d'interagir avec le geste, ainsi que du fait que la personne doit adapter une position adéquate afin d'être reconnu par les algorithmes.



FIG. 9 – Un eyetracker

L'eyetracker, tout comme la Kinect fonctionne sur la base d'un capteur optique infrarouge permettant de récupérer des informations relatives à la position du regard et aux mouvements effectués par les yeux. Il existe différents types d'eyetracker, tant au niveau de leurs aspects que de la technique qu'ils utilisent, et, de ce fait, des données qu'ils vont collecter. Certains modèles sont, par exemple, capables de détecter les variations du diamètre de la pupille de l'utilisateur. Ce qui peut s'avérer intéressant pour reconnaître des émotions : la peur ou la joie induisent des modifications physiologiques comme l'augmentation ou la diminution du diamètre de la pupille. La figure 9, présente l'EyeTracker développé par la société Tobii, une société suédoise spécialisée dans l'occulométrie. Ce dispositif se place en face de la personne, pour obtenir les meilleures informations possibles un calibrage est nécessaire. Celui-ci doit s'effectuer tant au niveau de la distance écran-eyetracker et eyetracker-sujet que de la taille du support sur lequel le sujet va porter son attention. Il existe d'autres dispositifs en forme de casque se fixant directement sur la tête, les capteurs se trouvant ainsi directement au niveau des yeux. Au delà des différences matérielles, il existe également différentes techniques telles que l'électro-oculographique (EOG), la technique du reflet cornéen ou de réflexion IR (IROG) ou encore la vidéo-oculographie (VOG), [Hari Singh, 2012].



FIG. 10 – Une surface tactile

Les surfaces tactiles sont, comme leur nom l'indique, des surfaces qui vont réagir lorsque des capteurs présents à la surface vont être activés. Au début de l'air du tactile en 1953, les capteurs étaient de types capacitifs et donc enregistrèrent une pression exercée sur la surface, ce qui les rendait peu précis. Les capteurs utilisés par la suite sont d'une tout autre nature. La société IBM en 1972 mettra sur le marché un dispositif tactile utilisant des infrarouges comme capteurs, ce dispositif est le fruit d'un projet nommé PLATO imaginé par un universitaire douze années auparavant. Par la suite, des évolutions sont apparues permettant ainsi à une surface de reconnaître plusieurs points de contact à la fois (écran multitouch) ou de ne réagir qu'avec certains matériaux tels que les doigts pour les téléphones portable ou les bornes SNCF ou les stylets pour d'autres périphériques.

Tout comme la souris, c'est Apple qui, avec la sortie l'iPhone 3, bouleversa à jamais le monde de l'interaction, en proposant un téléphone entièrement tactile, apportant ainsi une nouvelle vision de l'interaction. L'un des gros avantages du tactile est le fait de pouvoir être directement en contact avec ce que l'on interagit. Un autre avantage du tactile est ce que l'on appelle le retour haptique, ce dernier est défini comme étant une sensation physique ressentie par le toucher de l'écran comme si l'on venait d'appuyer sur un bouton mécanique. Bien que cette sensation physique soit pour la plupart du temps matérialisée par une « simple vibration » de l'écran, Senseg, une société finlandaise travaille actuellement sur la possibilité d'offrir la sensation à un utilisateur de toucher différentes textures à l'aide de modulation électrique à la surface des écrans tactiles⁶.

Il existe ainsi plusieurs technologies tactiles et chacune possède des avan-

6. <http://www.techniques-ingenieur.fr/actualite/1-avenir-du-tactile-peut-etre-dans-le-rendu-des-tes-68160/>

tages et des inconvénients⁷

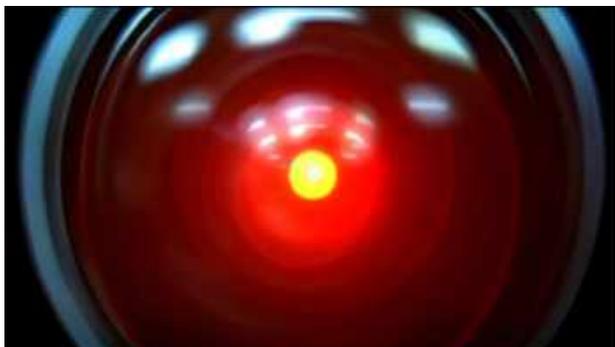


FIG. 11 – Le celebre HAL

Parmi ces nouveaux moyens d'interactions, le plus attendus de tous fut certainement la commande vocale. En effet, à l'inverse des autres dispositifs, la commande vocale possède cette particularité de ne pas être dépendante de son support, c'est-à-dire que le seul matériel nécessaire est un microphone. C'est pour cette raison qu'elle a suscitée tant d'intérêt, car la véritable barrière rencontrée par les développeurs était (et est toujours) le traitement du signal reçu. La commande vocale repose sur la langue naturelle et nous sommes donc confrontés à tous les problèmes du traitement automatique des langues (TAL). Bien que cette barrière fût longtemps un problème, à l'heure actuelle, les possibilités offertes par le progrès dans le domaine du traitement du langage naturel permettent d'embarquer ce nouveau moyen d'interaction dans certains dispositifs utilisés au quotidien. Pour s'en apercevoir il suffit simplement de regarder les services actuellement proposés par Siri (le dispositif de commande vocale implémenté dans les iPhone depuis 2011) ou Google Search (le dispositif intégré dans les terminaux Android). La reconnaissance vocale est donc possible dans toutes les conditions où la voix de l'utilisateur est audible, c'est-à-dire en marchant, dans un avion, dans le noir... La reconnaissance vocale, repose sur une des disciplines phare des sciences cognitives, le traitement automatique des langues (TAL).

⁷. http://www-igm.univ-mlv.fr/~dr/XPOSE2008/Les%20technologies%20tactiles/compo_pointage.html

1.2.3 Les Autres



FIG. 12 – Un casque EEG

Largement plébiscité quelques dizaines d'année auparavant dans les romans et films de science-fiction les plus célèbres, les interfaces cerveau-machine sont à l'heure actuelle une réalité et sont en pleine expansion. Le premier concept d'interface cerveau machine sera proposé en 1973 mais c'est en 1980 que les premiers essais cliniques sur l'homme seront menés par l'équipe de Thomas Elber⁸. Reposant sur notre connaissance du cerveau et de son fonctionnement, de tels dispositifs permettraient de se substituer à tout autre matériel.

Bien que très intrusifs à l'heure actuelle, les casques permette une interaction sans précédent puisqu'elle connecte directement l'homme à la machine, ce qui pourrait résoudre de nombre problème de handicap telle que l'utilisation d'Internet par la pensée ou toute autre activité connecté à un ordinateur ou une machine.



FIG. 13 – Le dispositif Leap Motion

8. <http://www.inserm.fr/thematiques/technologies-pour-la-sante/dossiers-d-information/interface-cerveau-machine>

Tout comme la Kinect, le Leap Motion⁹ est un dispositif de reconnaissance gestuelle fonctionnant à l'aide de capteurs infrarouges, beaucoup plus petit que cette dernière il offre une très grande précision. Contrairement à la Kinect, ce n'est pas tout le corps qui est reconnu mais uniquement les mains ou les objets se trouvant dans sa zone de reconnaissance, c'est à dire un cube virtuel de 60 cm de côté. De par sa précision et sa portabilité ainsi que son prix, ce dispositif remporte un franc succès. De nombreuses applications de contrôle d'ordinateur ou de manipulation d'objet en trois dimensions commencent ainsi à apparaître sur la plate forme d'application dédiés au dispositif qui après plusieurs mois d'attente sera enfin disponible au grand public dans les prochains mois.

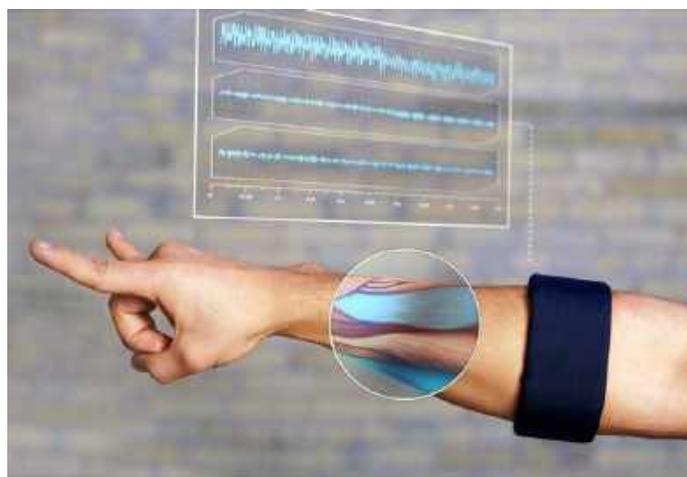


FIG. 14 – La technologie MYO

Alternative aux gants de données et présenté il y a quelques mois à l'état de prototype, le dispositif MYO¹⁰ est actuellement en cours de développement. Le dispositif se présente comme un bracelet se fixant sur l'avant bras de l'utilisateur afin de récolter des données physiologiques relatives aux mouvements effectués par l'utilisateur pour ensuite les associer à des actions logicielles. MYO est capable de détecter une rotation au niveau du poignet ou le simple fait de plier un doigt. Il dispose également de divers capteurs tels qu'un accéléromètre et un gyroscope qui permettent de détecter les déplacements effectués par le bras. Bien qu'il doive se porter, MYO est peu intrusif contrairement aux gants de données car le dispositif laisse libre les mains de l'utilisateur et n'apporte aucune contrainte de mobilité.

Les nouveaux moyens d'interaction nous permettent ainsi de nous détacher de la machine pour nous permettre de mieux y être connecté. En effet,

9. <https://www.leapmotion.com/>

10. <https://www.thalmic.com/myo/?autoplay=true>

bien que nous soyons toujours connectés à la machine pour pouvoir y interagir, les nouveaux moyens d'interaction offrent beaucoup plus de liberté aux utilisateurs tout en leur offrant une immersion sans précédent.

1.2.4 Conclusion

Il est intéressant de remarquer que même si la technologie cherche à se substituer au clavier et à la souris, l'avis des chercheurs n'est pas aussi tranché : Nicolas Roussel chercheur INRIA de Lille pense que la souris et le clavier ont encore de beaux jours devant eux¹¹. En effet, selon lui, il n'y a pas de moyen d'interaction meilleur qu'un autre, il soutient l'idée que chaque dispositif est approprié à certaines tâches. Ainsi, utiliser un ordinateur pour faire de la bureautique, surfer sur internet ou faire la comptabilité d'une entreprise ne sera jamais aussi efficace autrement qu'avec le couple clavier souris. L'argumentation de Nicolas Roussel suit le même raisonnement que Buxton [Buxton, 1986] : « Chaque dispositif possède des caractéristiques propres que le rendent extrêmement efficace pour un petit nombre de tâches, et peu approprié pour les autres ».

Ayant créé une application permettant d'utiliser une Kinect afin de naviguer dans Street View (applications sur laquelle nous reviendrons dans la dernière partie), il aurait été intéressant de comparer les performances des différents moyens d'interactions cités précédemment. Nous avons choisi de nous focaliser sur la Kinect, pour plusieurs raisons. Premièrement, nous avons décidé de ne pas négliger la contrainte temporelle et de nous concentrer sur un seul dispositif plutôt que de sous-estimer la charge de travail et de vouloir en faire de trop. Ensuite, participant à l'événement Renaissance 2013 avec d'autres binômes travaillant déjà sur l'eyetracker, présenter un autre dispositif d'interaction nous a semblé pertinent afin de présenter nos facultés d'adaptation et de représenter ainsi le potentiel de notre filière. De plus, après réflexion, nous nous sommes convaincu que l'eyetracker serait vraiment très peu adapté pour utiliser Street View. Enfin, vis-à-vis du tactile, nous avons choisi de l'écarter car à l'heure actuelle où environ 46% des français possèdent un smartphone, l'utilisation du tactile comme moyen d'interaction est chose courante, ce qui n'est pas le cas du geste.

Dans la partie suivante, nous allons donc nous intéresser à la composante fondamentale de la Kinect : le geste. Dans un premier temps nous allons nous interroger sur ce qu'est un geste, tant sur le plan général que dans le cadre d'une IHM puis dans un second temps nous nous intéresserons au geste dans le paradigme d'utilisation d'une Kinect.

11. <http://www.inria.fr/centre/lille/actualites/interaction-homme-machine>

2 L'interaction gestuelle

2.1 Qu'est ce qu'un geste ?

Qu'ils servent à appuyer le discours d'une personne qui parle ou à faire passer un message, les gestes font partie intégrante de notre communication [Kendon, 1972]. Un geste peut se caractériser comme un déplacement d'un membre ou d'une partie du corps sur au moins l'un des trois axes qui constituent l'espace à trois dimensions de notre environnement.

Cependant, dans un contexte de communication ou d'interaction homme-machine, le maintien d'une posture peut également être considéré comme un geste. Par exemple, le fait de tendre le bras à gauche peut engendrer une action brève et instantanée ou continue tant que la position est maintenue, ou encore indiquer une direction dans un contexte de communication. Il existe donc deux catégories de gestes, les gestes statiques et les gestes dynamiques. Cadoz, [Cadoz, 1994] catégorise les gestes selon leurs fonctions, distinguant les gestes épistémiques, ergotiques et sémiotiques :

- Les gestes épistémiques, servent à explorer le monde physique, par exemple lorsque nous touchons un mur pour en apprécier le relief.
- Les gestes ergotiques, nous permettent d'agir sur les objets présents dans ce monde, comme lorsque nous attrapons une pomme pour la manger.
- Les gestes sémiotiques, sont ceux qui nous permettent de communiquer avec nos semblables (gestes accompagnant la parole, langage des signes, etc.).

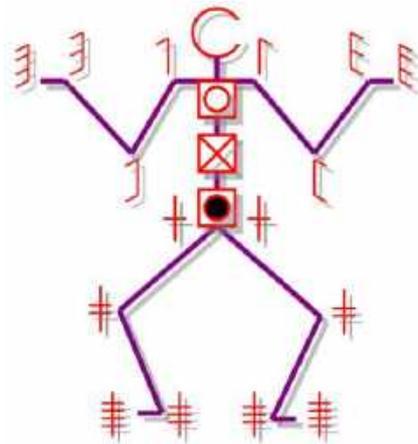


FIG. 15 – Codage des parties du corps selon la notation Laban

En 1928, Rudolf Laban met au point un système permettant de formaliser les gestes : la notation Laban ou encore Laban Kinetographique. Ce système s'appuie sur le squelette humain et segmente ce dernier en vingt

jointes représentés sur la figure 15. Pour coder un geste, la notation Laban utilise quatre vecteurs qui sont la force, le temps, l'espace et le poids. La notation Laban, bien que très efficace pour les chorégraphies, s'avère peu adaptée à certains mouvements fins et ne permet pas de donner du sens aux gestes. D'autres systèmes de transcription de gestes seront mis au point tel que le système de Stokoe qui en 1960 est le premier système adapté à la langue des signes.

Nous nous sommes donc inspiré du système Laban car nous n'avions pas besoin de de sémantique offerte par le deuxième système.

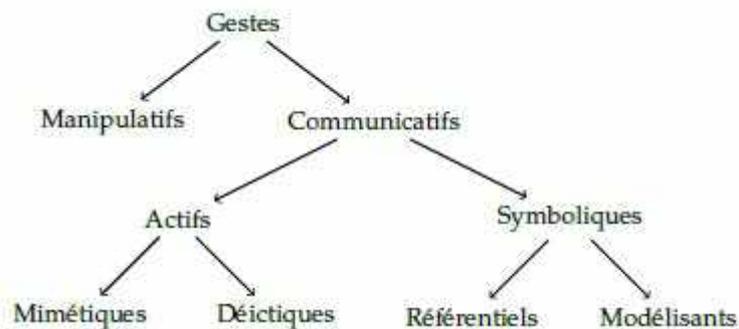


FIG. 16 – Taxonomie des gestes de Quek

Dans le cadre de l'interaction homme machine, la classification de Quek [Quek, 1994] visible sur la figure 16 permet de catégoriser les gestes et reprend sans les nommer les catégories de gestes proposées par Cadoz. Les gestes manipulatifs correspondent aux gestes épistémiques et ergotiques alors que les gestes communicatifs correspondent aux gestes sémiotiques de Cadoz. Les gestes symboliques font références aux gestes de la langue des signes par exemple car ils nécessitent un apprentissage pour en connaître la sémantique. Les gestes actifs restent les plus utilisés en interaction homme machine, en effet, les gestes déictiques représentent les gestes de pointages, permettant par exemple d'utiliser des mouvements de la main pour déplacer un pointeur dans une application. Quant aux gestes mimétiques, ils font références aux gestes mimant une action, dans le cas d'une interaction homme machine, le geste consistant à simuler avec ses pieds l'action de marche pour avancer dans Street View serait un geste mimétique.

2.2 Les gestes et la Kinect

Comme Baudel et Beaudouin-Lafon [Baudel & Beaudouin-Lafon, 1993] le soulignent, il y a un certain nombre d'avantages offerts par l'utilisation des gestes pour l'interaction :

- Une interaction naturelle : les gestes sont une forme naturelle de l'interaction et sont faciles à utiliser.
- Une interaction directe : le corps utilisé comme périphérique
- Simple et puissant, un geste simple peut être utilisé pour spécifier à la fois une commande ainsi que ses paramètres ;

Selon Gianni, [Gianni, 2007] il existe deux type d'interface gestuelles qui sont les interfaces utilisateurs perceptives dont le but « est de développer des interfaces conversationnelles à partir de ce qui est considéré comme un dialogue homme-homme naturel » et les Interface Utilisateur Multimodales ou « des gestes spécifiques sont utilisés pour passer des ordres dans un langage de commande ».

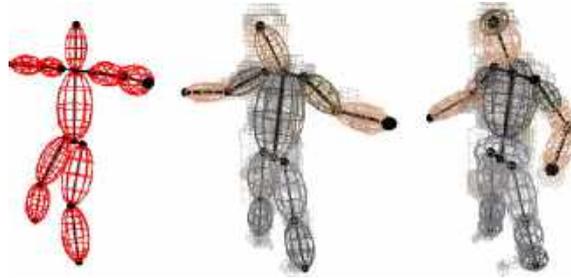


FIG. 17 – Les « blobs » en trois dimensions de Demirdjian

Contrairement à ce qui se faisait au début de la reconnaissance gestuelle, la Kinect offre une détection précise de l'utilisateur. En effet, il fut une époque où la reconnaissance des personnes se faisait via des formes géométriques (appelées « blobs »). Par exemple, les parties du corps étaient représentées sous formes de boîtes [Ju et al., 1996], sous forme de « blobs » en 2 dimensions [Wren et al., 1997], ou encore sous forme de « blobs » en trois dimensions [Demirdjian & Darrell, 2002] et [Nickel et al., 2004] comme présenté à la figure 17. Bien, qu'il permettaient une reconnaissance gestuelle, les « blobs » n'étaient pas en mesure d'offrir la précision d'une Kinect. Cependant, il est intéressant de remarquer sur la figure 17 qu'à l'intérieur des « blobs » se trouve le squelette de Laban.

Malgré une bonne détection de l'utilisateur, la Kinect possède des défauts typiques des périphériques de reconnaissance gestuelle via des flux vidéo. Parmi eux, le principal est sans doute le problème des gestes antagonistes dû au problème du repositionnement dans l'espace de l'utilisateur après avoir effectué un geste. Si l'on prend l'exemple de l'actions permettant de monter et de baisser la caméra dans un environnement virtuel afin d'en définir les gestes, nous serions tenté d'associer la première action au geste consistant à lever la main droite et la seconde au geste consistant à baisser cette même main. Bien que cela puisse être pratique et sensé pour un utilisateur, cela

s'avererra tout à fait contre productif. Pour cause, si l'utilisateur décide de monter la caméra par deux reprises car une seule fois n'est suffisante pour lui donner la vue souhaitée, il sera forcé de baisser sa main pour pouvoir la relever de nouveau. Or le fait de baisser la main induisant l'action de baisser la caméra, les actions effectuées seront : « monter caméra » « baisser caméra » « monter caméra » et non pas « monter caméra » « monter caméra » comme l'utilisateur l'aurait souhaité. Un autre défaut imputable à ce type de système est le manque de retour direct d'information qui fait d'eux des interface non haptiques c'est à dire qui ne permettent pas de donner de retour de toucher ou de retour de force, ce qui dans notre application pourrait donner l'impression à l'utilisateur de ne pas vraiment vraiment être en contact avec son environnement.

Il est important de ne pas confondre geste et action à réaliser : Une action peut aussi bien correspondre à un seul geste simple qu'à un ensemble de gestes qui seraient nécessaire pour la déclencher. L'action d'avancer à l'intérieur de Street View est l'affaire d'un seul geste cependant l'action de prendre l'intersection de droite demande une succession de gestes : pivoter vers la droite et se déplacer vers l'avant. Dans la section suivante, nous revenons sur le système de représentation d'un geste de la Kinect en le replaçant dans ce contexte.

3 La Kinect et Street View

3.1 L'application

3.1.1 Google Street

Google Maps est une application web mis en ligne par Google afin de pouvoir se localiser de manière rapide. Une carte virtuelle sur laquelle figure des informations (telles que le nom des villes, le nom des rues, les parcs, les cours d'eau) est proposée. L'interface permet de rechercher des lieux au sein de cette carte et de pouvoir se situer au sein d'une ville et -dans l'absolu- au sein du monde.

Street View est un service proposé par l'application Google Maps. Lancé en mai 2007 afin de compléter Google Maps et Google Earth, Street View permet de naviguer virtuellement dans les rues au sein des villes et sur les routes hors agglomérations au moyen de la technologie Immersive Media.¹² Il existe une API¹³ (Application Programming Interface, Interface de programmation en français) mise à la disposition des développeurs qui regroupe

12. <http://immersivemedia.com/>. Une voiture équipée de caméra circule dans les rues en prenant des images, celles-ci sont ensuite traitées par un logiciel et assemblées afin de donner une impression de continuité

13. <https://developers.google.com/maps/documentation/javascript/tutorial?hl=fr>

toutes les informations nécessaires si l'on veut utiliser Google Maps et ses services au sein d'une application web. Cette documentation concerne des fonctions en JavaScript et demande d'avoir des notions de programmation orientée objet.

Pour naviguer de panorama en panorama au sein de Street View plusieurs actions sont possibles, elles sont directement associées à la manipulation avec un clavier et/ou une souris. Il est donc important d'associer ces actions à leurs déclencheurs afin de garder le contexte dans lequel elles sont réalisées :

- clic gauche souris et maintenir appuyé : décaler la caméra vers le haut et vers le bas, tourner la caméra (gauche, droite, demi-tour).
- Actionner la molette : zoomer et dézoomer.
- clic gauche souris au sein d'un panorama : Changer de panorama.

Ces actions sont également déclenchées par des petits modules flottant mis à la disposition de l'utilisateur pour lui simplifier la navigation.

Street View a beau être un environnement discret, certaines actions donnent l'illusion d'un mouvement continu, cliquer et maintenir la souris pour orienter la caméra par exemple, le geste étant une unité de sens discrète il s'est posé la question de la manière de discrétiser l'action pour la commander par le geste. La solution que nous avons choisie s'apparente à celle des actions commandées par le clavier : un geste commande une action unique et les actions qui semblent continue doivent être résolues par une succession de gestes. Ainsi pivoter vers la droite de quatre-vingt-dix degrés demande une succession de geste associée à l'action de tourner à droite.

Pour permettre à l'utilisateur de naviguer au sein des panoramas il est donc nécessaire de savoir quels sont les panoramas environnant le panorama courant et de pouvoir situer le panorama et la position de caméra courante vis-à-vis de ceux adjacents. On peut visualiser l'ordonnancement de ces panoramas comme un graphe, chaque panorama possède les coordonnées de ses voisins au sein d'une liste. Chaque panorama ayant une valeur entre zéro et trois cent soixante pour le situer par rapport à la vue courante il n'aurait pas dû y avoir de problème. Seulement l'ordre des éléments contenus par la liste des voisins n'étant pas fixe d'un panorama à un autre il est nécessaire de recalculer chaque fois que l'on avance (ou que l'on recule) quel est le panorama devant (respectivement derrière). Si la valeur d'orientation dans la vue courante avait été comprise strictement entre zéro et trois cent soixante cela n'aurait encore pas posé de problème, seulement celle-ci n'est pas bornée. Quand on pivote vers la gauche cette valeur décroît (pouvant prendre des valeurs négatives) et s'incrément lorsque l'on pivote vers la droite. Il était donc nécessaire d'implémenter un algorithme qui permettrait de savoir quel panorama renvoyer quand on avance et quand on recule. Suite à des problèmes au sein de cet algorithme nous avons dû revenir à la solution d'émuler l'action sur les touches du clavier afin d'être opérationnel au cours de Renaissance. Il reste cependant certain qu'avec une meilleure compréhension de l'ordonnancement des panoramas ou une amélioration de

celle-ci l'algorithme reste la meilleure solution.

3.1.2 Les choix techniques

Kinect

Pour interagir avec la Kinect, deux SDKs étaient à notre disposition : Kinect for Windows et OpenNI. Kinect for Windows¹⁴ présente l'avantage d'être le SDK officiel fourni par Microsoft et donc de posséder une documentation officielle sur laquelle s'appuyer, raison principale pour laquelle nous l'utilisons.

Dans le cadre d'une interaction avec la Kinect plusieurs langages de programmation se sont révélés judicieux. En effet, avec le SDK de Kinect for Windows des exemples de programmes sont fournis¹⁵ et trois langages se démarquent au sein de ces logiciels : Visual Basic, C++ et C#. Le C# ayant une documentation plus fournie au niveau des classes faisant le lien avec la Kinect et les exemples programmés dans ce langage étant plus élaborés que les autres, nous avons opté pour ce langage pour programmer notre application.

Lors de notre état de l'art, trois sources d'informations principales sont ressorties : le forum officiel de Microsoft¹⁶, une application ressemblant à ce que devrait être la notre appelée Kinect master¹⁷ et le livre de Catuthe [Catuthe, 2012] Programming with the Kinect for Windows Software Development Kit. Elles nous ont permis d'avoir une base solide sur laquelle nous appuyer aussi bien au niveau de l'interface graphique, utilisation du xaml¹⁸, que de la liaison avec la Kinect.

Le but de l'application étant de naviguer au sein de Street View, il a été nécessaire d'émuler un navigateur. Contrairement à Kinect Master qui simule simplement des boutons du clavier, pour plus de contrôle sur Maps et Street view et grâce à l'API mise à notre disposition par Google, nous faisons appel à des fonctions JavaScript stockées dans les pages html appelées par le navigateur.

Comme nous l'avons vu précédemment la Kinect ne reconnaît pas directement le corps mais reconstruit un squelette sommaire à partir des formes et des mouvements qu'elle reconnaît. Au sein du SDK une classe représentant celui-ci existe : la classe Skeleton. Il existe une méthode qui permet de récupérer toutes les instances des squelettes suivit par la Kinect. Il suffit ensuite de créer des détecteurs -qui hérite de la classe GestureDetector- qui, à partir des mouvements des joints, lèvent des événements lorsque le mouvement

14. <http://www.microsoft.com/en-us/kinectforwindows/>

15. <http://www.microsoft.com/en-us/kinectforwindows/develop/developer-downloads.aspx>

16. <http://msdn.microsoft.com/en-us/library/hh855347.aspx>

17. <https://github.com/syoummer/KinectStreetView>

18. eXtensible Application Markup Language : langage qui permet la définition des données structurées <http://www.xaml.fr/>

correspondant aux gestes prédéfinis sont détectés. Le programme principal déclenche de manière réactive les actions associées à ces évènements qu'il récupère par le biais de gestionnaire d'évènement.

3.1.3 Description des classes

Ayant défini les outils techniques nous allons expliquer à présent le fonctionnement des classes C# et des fonctions JavaScript que nous avons élaboré à l'intérieur de notre application.

Classes réalisées en C#, 869 lignes

- Candidat, classe reprenant un candidat avec son temps et son parcours, 30 lignes
- CandidatComparer, classe qui permet de comparer deux candidats par rapport à leur temps respectif, 28 lignes
- Classement, classe qui sert à enregistrer les candidats dans un fichier texte, à trier selon leur performance et à les charger depuis le fichier, 52 lignes
- StreetViewGestureDetector3D, hérité de GestureDetector cette classe sert à analyser le flux de donnée renvoyé par la Kinect afin de lever un évènement différent pour chaque geste reconnu, 254 lignes
- MainWindow, classe principale, elle gère les détecteurs qui suivent les joints du squelette reconstruit, les actions sur les navigateurs lorsque les évènements correspondant sont levés, elle charge le classement à partir du fichier associé. Elle fait aussi le lien avec la Kinect , l'initialisant et la maintenant active. C'est aussi elle qui déclenche les évènements associés à la fenêtre XAML, 450 lignes
- Le XAML associé à MainWindow, fenêtre principale, elle affiche les navigateurs afin d'interagir avec Street View et Google Maps, une fenêtre affichant le résultat filmé par la Kinect et la reconstruction des squelettes lorsqu'ils sont détectés et les boutons pour interagir avec l'application, 55 lignes

Pages html réalisées comprenant des fonctions JavaScript, 340 lignes

- j1_map.html, 94 lignes, comprenant les fonctions :
 - initialize(latitude,longitude), qui charge Google Maps au sein de la page et initialise aux coordonnées latitude, longitude.
 - setPosition(latitude,longitude), qui bouge le centre de la carte aux coordonnées passées.
 - dessiner_chemin(position), qui met à jour la liste correspondant au chemin pris par l'utilisateur lui rajoutant les coordonnées position et qui retrace le chemin.
 - affiche_texte(texte), qui permet d'afficher un texte dans un cadre en haut de la carte.
- j1_street_view.html, 153 lignes, comprenant les fonctions :

- initialize(), qui charge Street View au sein de la page, les coordonnées du panorama de départ sont contenues dans la page.
- avancer(), qui fait avancer d'un panorama en se basant sur l'orientation courante de la vue.
- reculer(), qui fait reculer d'un panorama en se basant sur l'orientation courante de la vue.
- latitude(), qui récupère la latitude du panorama actuel.
- longitude(), qui récupère la longitude du panorama actuel.
- gauche(rapidite), qui modifie l'orientation courante de la vue vers la gauche en fonction de la rapidité demandée.
- droite(rapidite) , qui modifie l'orientation courante de la vue vers la droite en fonction de la rapidité demandée.
- monter(rapidite) , qui modifie l'orientation courante de la vue vers le haut en fonction de la rapidité demandée.
- descendre(rapidite) , qui modifie l'orientation courante de la vue vers le bas en fonction de la rapidité demandée.
- getZoom(), qui récupère le zoom courant.
- zoomAvant(forceZoom), qui augmente le zoom courant en fonction de la force demandée.
- zoomArriere(forceZoom), qui diminue le zoom courant en fonction de la force demandée.
- map_centrale.html, 67 lignes, comprenant les fonctions :
 - initialize(latitude,longitude), qui charge Google Maps au sein de la page et initialise aux coordonnées latitude, longitude.
 - setPosition(latitude,longitude), qui bouge le centre de la carte aux coordonnées passées.
 - dessiner_chemin(position), qui met à jour la liste correspondant au chemin pris par l'utilisateur lui rajoutant les coordonnées position et qui retrace le chemin.
 - affiche_texte(texte), qui permet d'afficher un texte dans un cadre en haut de la carte.
- finale.html, 26 lignes, comprenant la fonction :
 - affiche_texte(texte), qui permet d'afficher un texte au sein de l'image choisie pour annoncer la fin de la course.

3.2 La collecte des gestes

Ayant choisit d'utiliser le canal gestuel pour interagir avec Street View via le dispositif Kinect, nous avons choisi dans un premier temps de collecter des gestes auprès de sujets. Les gestes ainsi relevés nous permettrons d'obtenir un panel de gestes sur lesquelles nous pourrions formuler des hypothèses. Comme l'explique [Gianni, 2007] il existe de nombreux systèmes de notations de gestes, comme labanotation ou HamNoSys et SignWriting pour la langue des signes... Cependant, bien que très efficace, la notation Laban est

trop complète (pour les gestes que nous souhaitons récupérer) et relativement difficile à prendre en main. Nous nous sommes donc inspirés de son fonctionnement pour créer notre propre système de notation. Notre système se compose de grilles permettant de matérialiser à la fois l'espace et le temps. Tout comme la notation Laban, chaque joint du corps est matérialisé par un symbole.

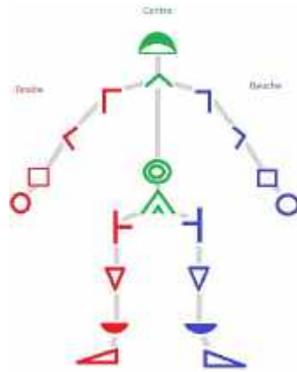


FIG. 18 – Symbolique des joints

La figure 18 représente le codage que nous avons utilisé, codage qui était plus facile à apprendre d'une part mais surtout qui permet d'utiliser des formes distinctes pour chaque joint permettant ainsi de relever des gestes plus rapidement. Ce squelette est identique à celui de Laban puisqu'il utilise les mêmes joints, qui sont les joints utilisés par la Kinect.

	Face	Profil	Dessus
T=0			
T=1			
T=2			

FIG. 19 – Grille de relevé de geste

La figure 19 correspond aux grilles de relevé, chaque colonne correspond

à un point de vue illustré par la figure suivante et chaque ligne correspond à un pas de temps. Le choix du nombre de temps pour définir le geste s'est posé : il fallait un nombre suffisamment grand pour discriminer facilement les gestes mais pas trop pour faciliter la récolte. Trois pas de temps nous a paru un bon compromis entre ces deux objectifs. En plus des neuf grilles permettant le relevé des geste, un nom peut également lui être donné, cela permet de gagner du temps lors de la collecte si un geste déjà codé est de nouveau effectué par un autre sujet.

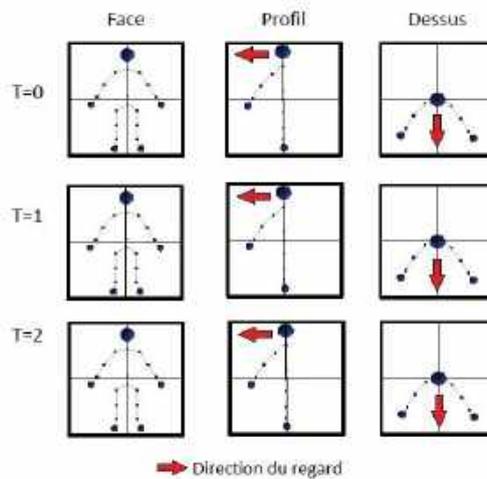


FIG. 20 – Point de vue de la grille

La figure représente la position de la personne telle qu'elle doit être observée par l'opérateur qui effectue le relevé du geste effectué.

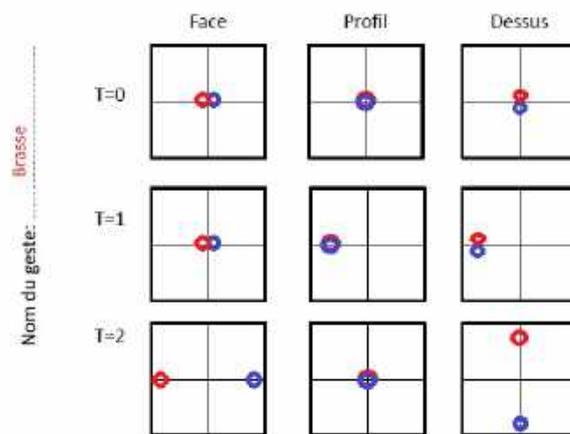


FIG. 21 – Codage du geste "brasse"

Si l'on prends le geste qui représente un mouvement de brasse en natation, fait uniquement avec les bras, nous obtenons la figure 21 ci-dessus.

Ce système de notation permet donc un codage direct des gestes, c'est-à-dire que nous n'avons pas besoin d'avoir recours à un système d'enregistrement vidéo qui aurait nécessité un visionnage ultérieur pour coder les gestes.

Au total, vingt-cinq personnes dans la tranche 18-25 ans ont été interrogées avec unique consigne : "quels sont les gestes que vous voudriez utiliser pour [action] " action=avancer, reculer, aller à gauche, à droite... Suite à cela, plusieurs questions ont été posés aux sujets sur leur appréciation de ce type d'interaction, ces dernières ainsi que les réponses étant disponibles en annexe dans la section "relevé de gestes et questionnaire".

Nous avons par la suite, répartis les gestes par leur fréquence d'apparition, sans les dissocier de leur action associée¹⁹, afin d'en dégager deux catégories :

- Les « tops gestes » ; gestes les plus fréquents.
- Les « flops gestes » : gestes les moins fréquents.

Un troisième groupe de gestes appelés prototypiques a été défini par nos soins. Ils sont donc pensés pour être efficaces et agréables dans le sens où nous avons une connaissance du moyen d'interaction et un recul inaccessible à un utilisateur lambda.

Action/Catégorie	Top	Flop
Avancer	Tirer main droite devant	Ouverture ascenseur
Reculer	Pousser main droite devant	Fermeture ascenseur
Aller à droite	Balayage extérieur main droite	Nage main droite sans passer devant
Aller à gauche	Balayage extérieur main gauche	Nage main droite sans passer devant
Regarder en haut	Lever tête	Monte haltère main droite
Regarder en bas	Baisser tête	Rejoindre mains vers le bas bras tendus
Zoom avant	Ouverture ascenseur	Cercle main gauche bras tendu horaire
Zoom arrière	Fermeture ascenseur	Cercle main gauche bras tendu anti-horaire
Aller dans Street View	Diagonale main droite montante vers gauche	Moulinette deux mains horaire
Aller dans Google Map	Diagonale main droite montante vers gauche	Moulinette deux mains anti-horaire

TAB. 1 – Gestes récolter pour naviguer dans street View

3.3 Une course folle

Afin de mesurer les performances de la Kinect pour naviguer dans Street View, nous avons créé une course au sein de notre application. Un point de départ ainsi qu'une ligne d'arrivée ont donc été implanté dans Street View,

¹⁹. Pour savoir si le geste de la brasse fait partie des Tops ou des Flops pour l'action avancer, on compte son nombre d'apparition pour cette action parmi les gestes proposés.

le but de la course étant de réaliser le meilleur temps possible.

Notre objectif initial était de comparer les performances des moyen d'interaction disponible dans le projet D-Min : Eyetracker, Kinect, couple clavier/souris et surface tactile. Nous avons ensuite réfléchis à ce que cela requièrerait vis-à-vis de la Kinect et une autre perspective s'est imposée à nous : avant de comparer l'efficience du moyen d'interaction, il convenait de définir des gestes pour se déplacer au sein de Street View.

Ces gestes auraient pu être définis de manière arbitraire mais nous n'aurions pas alors été en mesure de savoir si les différences d'efficacité dépendaient du choix des gestes ou de la Kinect pour elle-même. La course nous paraissait donc un outil qui nous permettrait d'évaluer nos panels de gestes entre eux. Nous avons donc voulu comparer les performances des trois groupes de geste (tops, flops et prototypiques). Nos hypothèses étaient les suivantes :

- Les gestes flops ne permettent pas d'obtenir de bons score à la course.
- Les gestes top et les gestes prototypiques permettent d'obtenir des scores identiques.
- Les gestes flops sont moins agréables à utiliser que les gestes prototypiques ou les gestes flops.

Cependant, le cadre offert par Renaissance 2013 ne nous permettait pas d'obtenir les conditions requises pour mesurer correctement les performances des différents groupes. Ce point fera d'ailleurs l'objet d'une discussion dans la partie suivante.

Nous avons concentré notre réflexion sur la problématique dont nous avons parlé dans l'introduction, qui est de savoir si la Kinect est un moyen d'interaction adaptée pour naviguer dans Street View.

Pour ce faire, nous nous sommes donc servis de questionnaires, que nous avons fait passer aux personnes auprès desquelles nous avons récolté des gestes puis aux utilisateurs de notre application.

Les questions posés dont les réponses sont disponibles en annexe sont les suivantes :

- Question : Pour des dispositifs d'interaction avec le geste, pensez vous que les gestes doivent être définis au préalable par le « constructeur » (Kinect, Leap Motion, MYO...)?
- Question 2 : La Kinect (les gestes) vous parait elle le moyen le plus efficace pour utiliser Google Street ?
- Question 3 : La Kinect (les gestes) vous parait elle le moyen le plus agréable pour utiliser Google Street ?
- Question 4 : Quel moyen d'interaction vous parait il le plus adapté pour utiliser Google Street ?

Nous expliciterons dans la partie suivante les résultats que nous avons obtenus ainsi que les conclusions que nous avons pu en dégager.

3.4 Résultats et analyse

Questions/reponses	OUI	NON
Prédefinitions des gestes	15	10
Kinect plus efficace	11	14
Kinect plus agréable	18	7

TAB. 2 – Réponses des personnes interrogées lors de la récolte des gestes

Questions/reponses	OUI	NON
Prédefinitions des gestes	18	7
Kinect plus efficace	10	15
Kinect plus agréable	14	11

TAB. 3 – Réponses des personnes interroger sur le stand de Renaissance 2013

Afin de voir si l'utilisation de la Kinect pouvait influencer l'avis des personnes quand à sa pertinence dans l'utilisation de Street View, nous avons effectué un test statistique comparant deux groupes. Le premier groupe étant composé de personnes n'ayant pas utilisé notre application²⁰ et le second étant constitué des personnes que nous avons interrogées lors Renaissance 2013 après que ces dernières aient utilisé notre application.

Le test du Chi^2 de Pearson a été effectué avec le logiciel SPSS afin comparer le nombre de réponses positives à chaque question avant et après l'utilisation de la Kinect et de Street View²¹.

Les Chi^2 obtenus ne se sont pas révélés significatifs, ce qui implique que statistiquement, nous ne pouvons pas dire que les résultats obtenus ne soient pas dû au "hasard".

Nous pouvons tout de même constater certaines tendances comme le fait que les personnes sont en majorités d'accord pour que les gestes soient pré-définis par les concepteurs. De plus, en regardant les justifications que nous ont fournis les sujets, il apparait que les personnes souhaitent tout de même pouvoir avoir la possibilité de modifier les gestes pour les rendre personnels.

20. ce sont ces même personnes que nous avons interrogées pour la récolte des gestes

21. (Question 1 : $Chi^2=0,802$; $p=0,276$, Question 2 : $Chi^2=0,082$; $p=0,5$, Question 3 : $Chi^2=0,802$; $p=0,276$)

Réponses au questionnaire

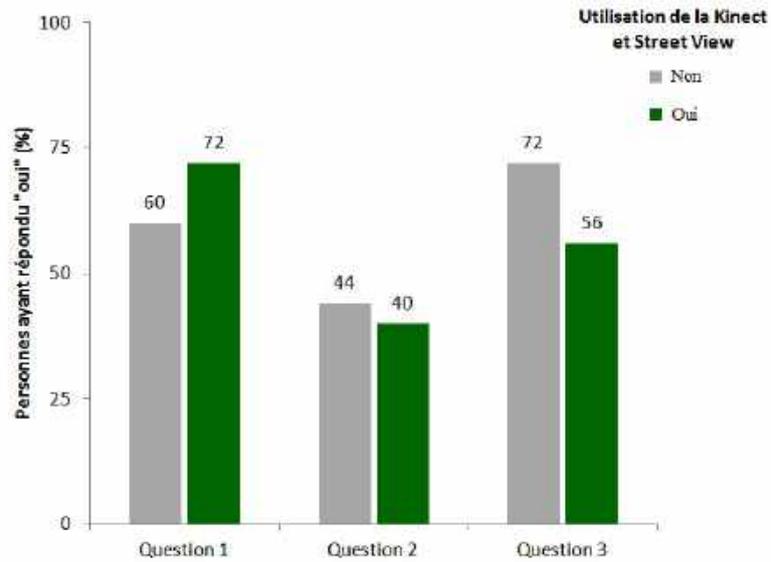


FIG. 22 – Résultats questionnaires

Le graphique de la figure 22 représente le réponses "oui" obtenues aux 3 premières question du questionnaire.

Moyens d'interaction	Nombre de choix
Clavier souris	7
Souris	4
Kinect	5
Tactile	3
Joystick	2
Manette	2
Vocal	1
Casque EEG	1
Clavier	0
Total	25

TAB. 4 – Réponses du 1er groupe

Moyen d'interaction	Nombre de voix
Clavier souris	9
Souris	6
Clavier	1
Kinect	5
Tactile	3
Leap Motion	1
Total	25

TAB. 5 – Réponses du 2ème groupe

Les deux tableaux ci-dessus, résument les réponses obtenues à la quatrième questions concernant les deux groupes définis précédemment.

Un test du Chi^2 à également été effectué afin de comparer les moyens d'interactions proposés par les sujet comme étant les moyens les plus adaptés pour naviguer dans Street View. Les détails de ce test sont disponibles en annexe dans la partie intitulée "analyse Chi^2 ".

Le Chi^2 n'est pas significatif, nous ne pas exclure le fait que les résultats obtenue puissent être du au hasard. Cependant nous pouvons constater que les personnes on tendance à proposer plus le plus souvent le clavier et la souris.

Le fait que les résultats ne soient pas significatifs ne signifie pas qu'il n'y que l'on ne peut rien démontrer mais que les moyens mis en œuvre n'ont peut être pas été suffisant.

Afin de résoudre le problème il faudrait augmenter notre puissance statistique. Pour ce faire il faudrait questionner de nouveaux utilisateurs et consigner à nouveau leurs avis.

Lorsque l'on pose les questions il faudrait s'assurer que l'utilisateur juge l'interaction avec la Kinect elle-même et pas au travers de l'application qui lui sert de support. Afin de résoudre ce problème nous pourrions utiliser la technique du magicien d'Oz. Cette dernière consiste à présenter une application à l'utilisateur mais ce n'est pas lui qui agit directement sur celle-ci : un compère connaissant l'application simule les actions associées à ce que fait l'utilisateur afin que celui-ci ne se concentre pas sur ses difficultés à réaliser une tâche.

La récolte des données ayant été opérée au sein d'un stand de démonstration, cela permettait d'avoir facilement un panel assez diversifié et aurait pu nous permettre d'agrandir celui-ci. Cependant le contexte public comporte des inconvénient majeurs tels que le bruit, le fait de passer devant les autres, le manque de temps de certaines personnes, le fait que les utilisateurs ne se sont pas réellement engagés à pratiquer l'expérience mais étaient là pour une démonstration etc. Autant de biais qui pourraient influencer les résultats, il

serait donc intéressant de faire repasser ces tests dans un environnement plus neutre et propice à ce genre d'exercice.

Bien que nos résultats ne nous aient pas permis de déterminer si effectivement la Kinect est adaptée ou non à l'utilisation de Street View, le fait que les personnes n'aient pas les mêmes attentes vis-à-vis de l'application peut être un obstacle pour répondre à cette question. Ce qui rejoint la conclusion de Nicolas Russel : "Il n'y a pas de moyen d'interaction parfait, tout dépend du contexte (performances, exploration etc...) dans lequel cette interaction est réalisée".

Bilan personnel

Ce projet, fut tout d'abord une expérience très enrichissante à de nombreux niveaux en plus de nous avoir donné une bonne approche du métier de chercheur. Nous avons donc dû faire preuve de rigueur tout au long de ce projet, nous avons ainsi dû apprendre à gérer le temps qui nous était imparti afin de l'exploiter de la manière la plus efficace qui soit. Pour ce faire, nous nous sommes fixés à l'aide de notre directeur de projet tutoré des rendez-vous réguliers avec à chaque fois des objectifs qui devaient être réalisés. Cela nous a permis de voir notre projet se construire peu à peu tout en écartant le risque de partir dans de mauvaises directions. Notre projet nous a également offert la possibilité de travailler en équipe, car ce dernier était en relation avec deux autres projets qui tout comme le notre ont eu la chance de pouvoir être présent à l'exposition Nancy Renaissance 2013. L'enthousiasme et l'intérêt des utilisateurs qui ont utilisé notre application fut très appréciable et encourageante. La confrontation avec d'autres binômes nous a permis d'obtenir des points de vue et des discussions intéressantes tant sur le plan personnel que sur le plan professionnel. Ce projet fut d'autant plus intéressant qu'il touche aux nouvelles technologies qui sont en pleine expansion. Nous avons donc dû prendre du recul par rapport à ces technologies afin de d'aborder notre problématique d'un point de vue scientifique et non pas d'un point de vue « consommateur ». Nous avons ainsi, grâce au travail que nous avons fourni, gagné en maturité tant au niveau de la rédaction d'un rapport que de la gestion d'un projet en groupe. Nous tenons particulièrement à remercier notre directeur de projet tutoré pour nous avoir appris à nous poser les bonnes questions, ainsi que toutes les techniques que nous avons apprises pour mener à bien notre projet. Nous remercions également les différentes personnes que nous avons pu rencontrer au LORIA, qui nous ont apporté conseils et intérêt sur notre travail.

Références

- [Baudel & Beaudouin-Lafon, 1993] Baudel, T. & Beaudouin-Lafon, M. (1993). *Communications of the ACM* 36, 28–35.
- [Buxton, 1986] Buxton, W. (1986). In *User Centered System Design : New Perspectives on Human-Computer Interaction*, (Norman, D. A. & Draper, S. W., eds), pp. 319–337. Erlbaum Hillsdale, NJ.
- [Cadoz, 1994] Cadoz, C. (1994). *Technique et Science de l'Information* 13(1), 31–61.
- [Catuhe, 2012] Catuhe, D. (2012). *Programming with the Kinect for Windows Software Development Kit. Developer reference*, Microsoft GmbH.
- [Demirdjian & Darrell, 2002] Demirdjian, D. & Darrell, T. (2002). In *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces ICMI '02* pp. 267–, IEEE Computer Society, Washington, DC, USA.
- [Gianni, 2007] Gianni, F. (2007). *Suivi de parties du corps pour l'interprétation de gestes de communication à partir de séquence monoculaire*. PhD thesis, Université Paul Sabatier Toulouse III.
- [Hari Singh, 2012] Hari Singh, D. J. S. (2012). *International Journal of Scientific and Research Publications* 2.
- [Ju et al., 1996] Ju, S. X., Black, M. J., & Yacoob, Y. (1996). In *2nd Int. Conf. on Automatic Face- and Gesture-Recognition* pp. 38–44,, Killington, Vermont.
- [Kendon, 1972] Kendon, A. (1972). *Some relationships between body motion and speech* pp. 177–216. Pergamon Press.
- [Nickel et al., 2004] Nickel, K., Seemann, E. & Stiefelhagen, R. (2004). In *Proceedings of the Sixth IEEE international conference on Automatic face and gesture recognition, FGR' 04*, pp. 565–570, IEEE Computer Society, Washington, DC, USA.
- [Quek, 1994] Quek, F. K. H. (1994). In *Virtual Reality Software and Technology : Proc. of the VRST'94 Conference*, (Singh, G., Feiner, S. K. & Thalmann, D., eds), pp. 17–31. World Scientific London.
- [Wren et al., 1997] Wren, C. R., Azarbayejani, A., Darrell, T. & Pentland, A. (1997). *IEEE Trans. Pattern Anal. Mach. Intell.* 19, 780–785.

.1 Analyse chi2

Formule utilisée pour calculer le Chi^2 :

$$\left(\frac{(Observe - Theorique)^2}{Theorique} \right)$$

Nombre de degré de liberté = (nombre de ligne- 1)*(nombre de colonnes – 1) = (2-1)*(10-1)=9 La valeur du Chi^2 donné par une table avec neuf degrés de liberté pour probabilité de 0.95 est 16.919

avant= personne interrogées pour les gestes elles n'ont donc pas utiliser notre application

après = personnes interrogées à Renaissance, elles ont donc utilisé notre application

Périphérique/groupe	Clavier et souris	Souris	Kinect	Tactile	Joypad	Manette	Clavier	Casque EEG	Leap Motion	Vocal	Total
Avant	7	4	5	3	2	2	0	1	0	1	25
Après	9	6	5	3	0	0	1	0	1	0	25
Total	16	10	10	6	2	2	1	1	1	1	50

TAB. 6 – Réponse question 4

Périphérique/groupe	Clavier et souris	Souris	Kinect	Tactile	Joypad	Manette	Clavier	Casque EEG	Leap Motion	Vocal	Total
Avant	0.125	0.2	0	0	1.5	0.5	0.25	0.25	0.25	0.25	2.325
Après	0.125	0.2	0	0	1.5	0.5	0.25	0.25	0.25	0.25	2.325

TAB. 7 – Contribution Chi^2 question 4

Le valeur du Chi^2 est donc de 4.65, la valeur théorique du Chi^2 telle que l'on peut la trouver dans une table pour neuf degrés de liberté et une probabilité d'erreur de 0.05 est de 16.919. La valeur du Chi^2 obtenue étant inférieur, l'hypothèse nulle ne peut donc pas être rejetée.

.2 Relevé de gestes et questionnaire

Action: Avancer

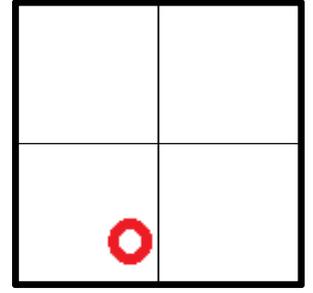
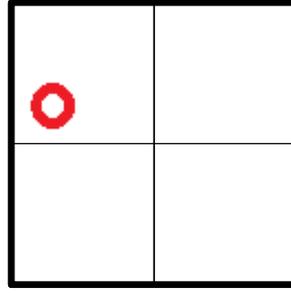
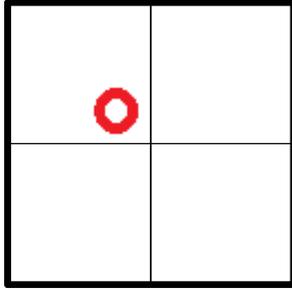
Nom du geste: Tirer main droite devant

Face

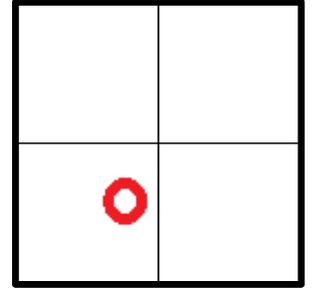
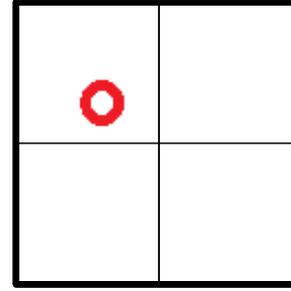
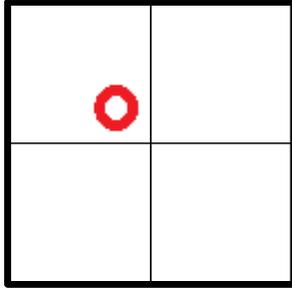
Profil

Dessus

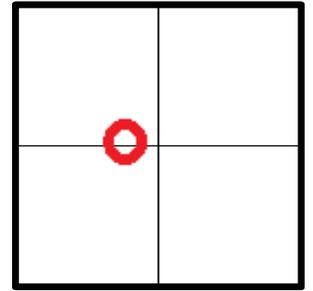
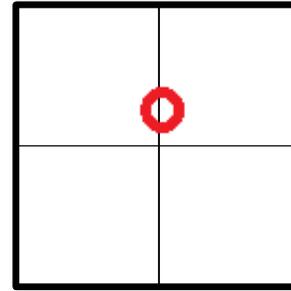
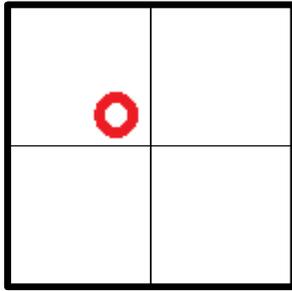
T=0



T=1



T=2



Action: Reculer

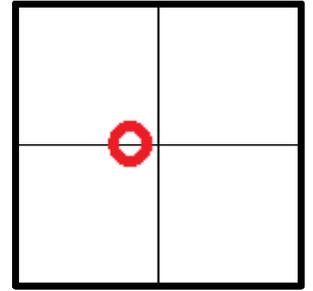
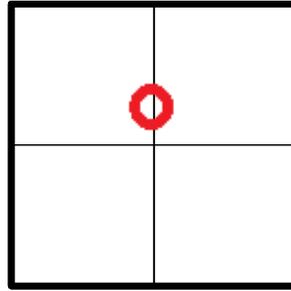
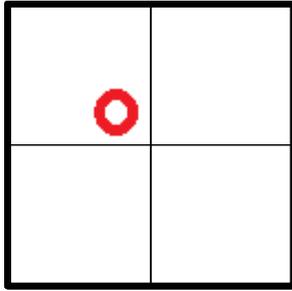
Nom du geste: Pousse main droite devant

Face

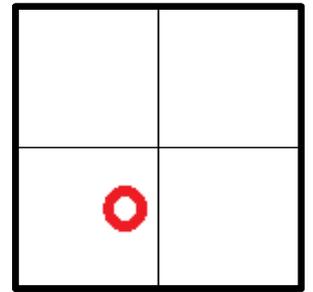
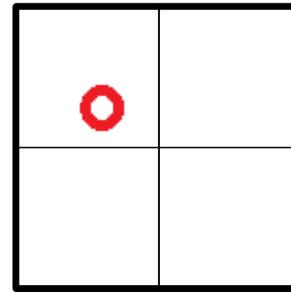
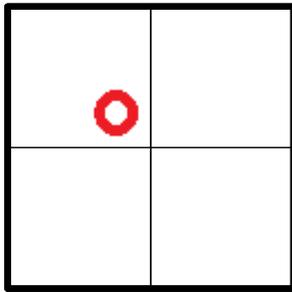
Profil

Dessus

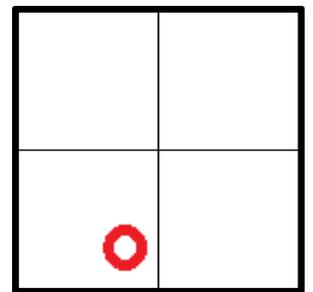
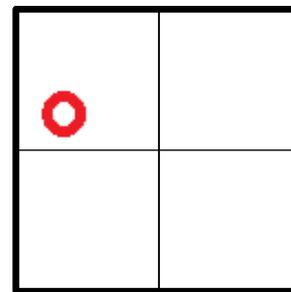
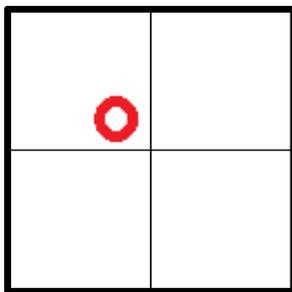
T=0



T=1



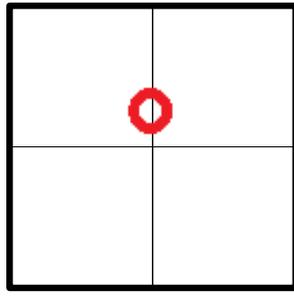
T=2



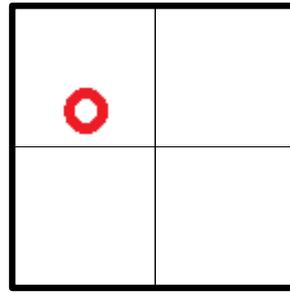
Action: Aller à droite

Nom du geste: Balayage extérieur main droite

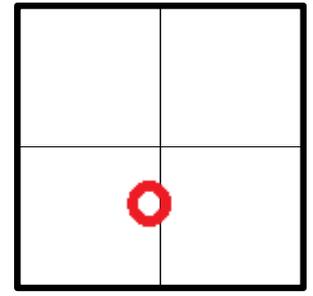
T=0



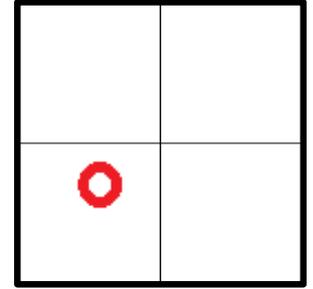
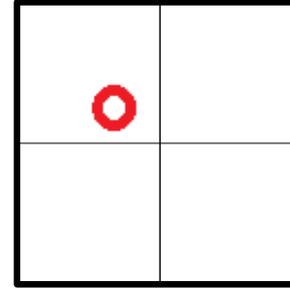
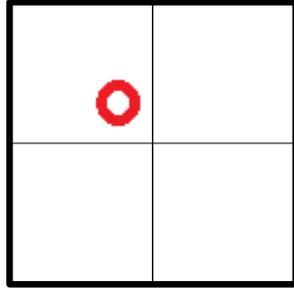
Profil



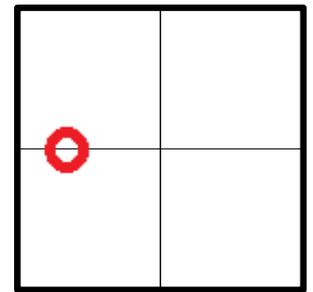
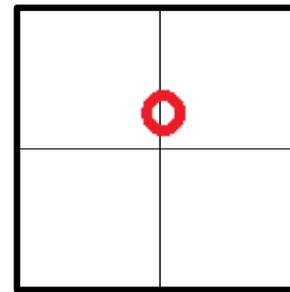
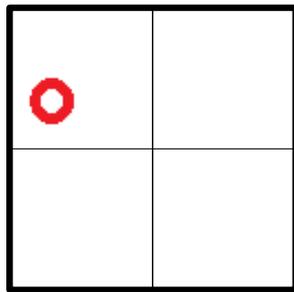
Dessus



T=1



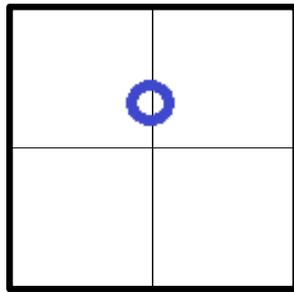
T=2



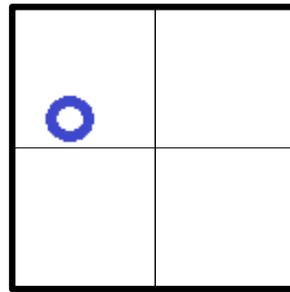
Action: Aller à gauche

Nom du geste: Balayage extérieur main gauche

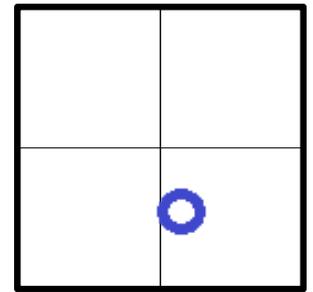
T=0



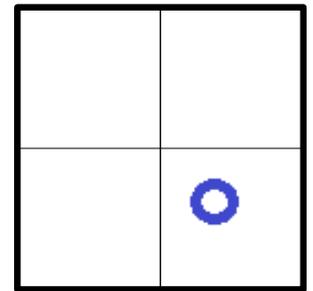
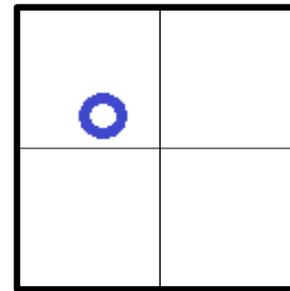
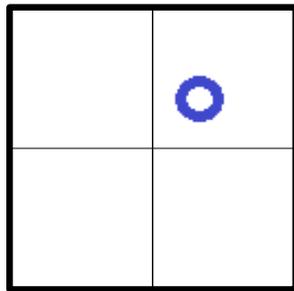
Profil



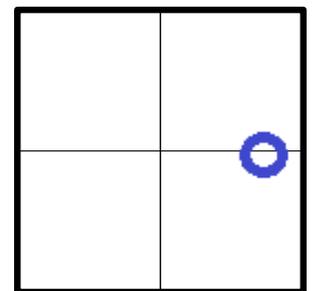
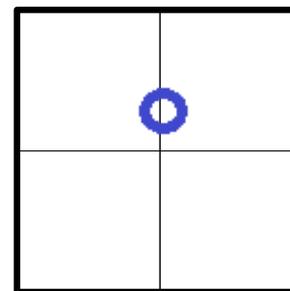
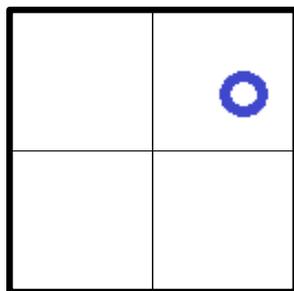
Dessus



T=1



T=2



Face

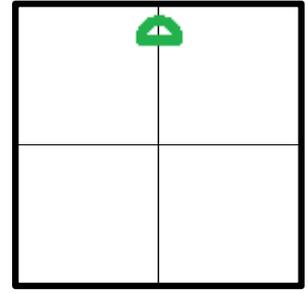
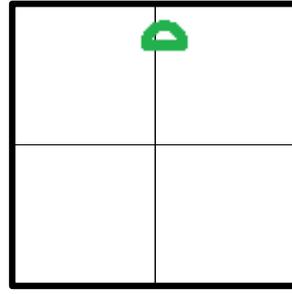
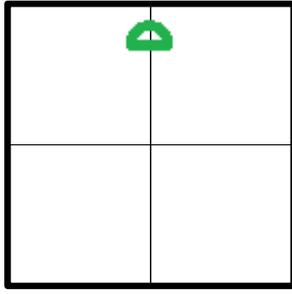
Profil

Dessus

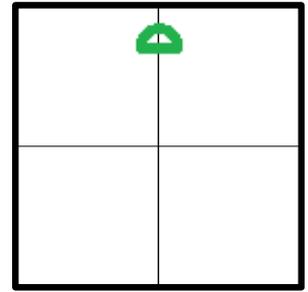
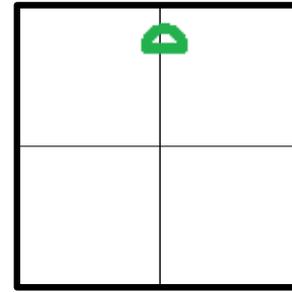
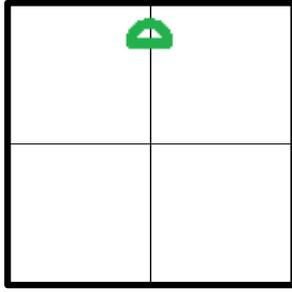
Action: Regarder en haut

Nom du geste: Lever tête

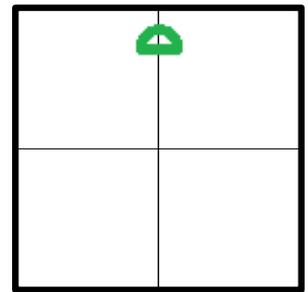
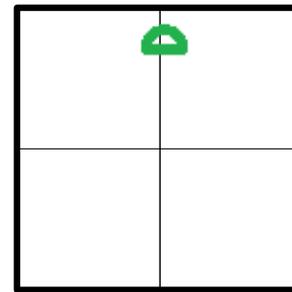
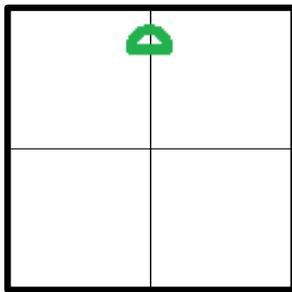
T=0



T=1



T=2



Face

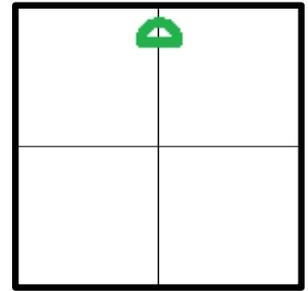
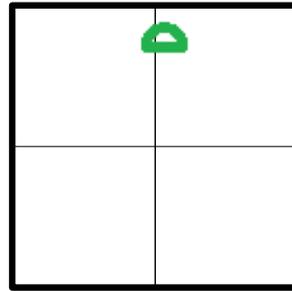
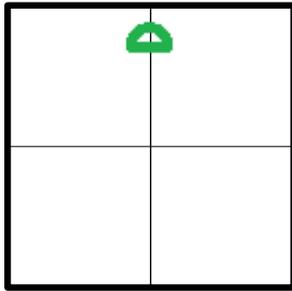
Profil

Dessus

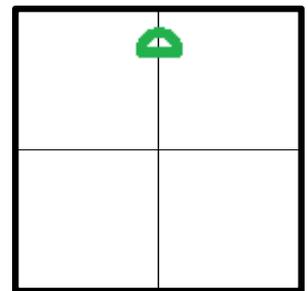
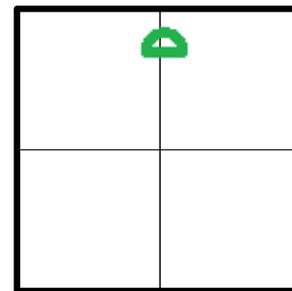
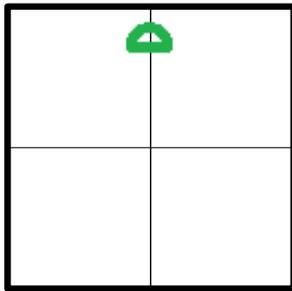
Action: Regarder en bas

Nom du geste: Baisser tête

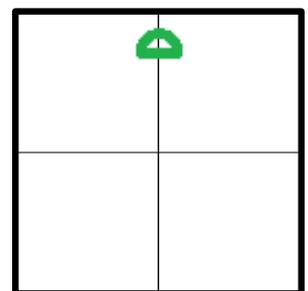
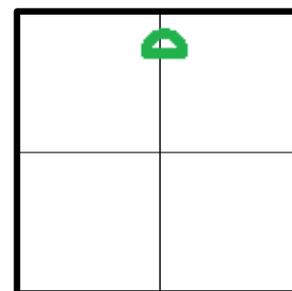
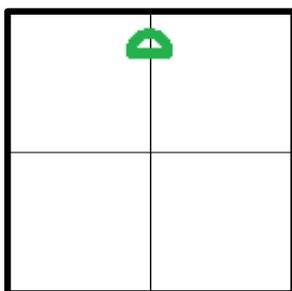
T=0



T=1



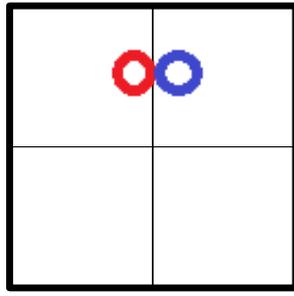
T=2



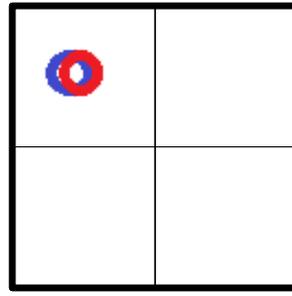
Action: Zoom avant

Nom du geste: Ouverture ascenseur

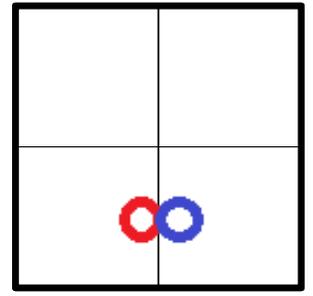
T=0



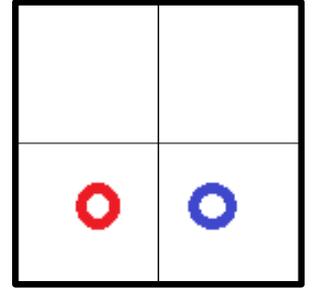
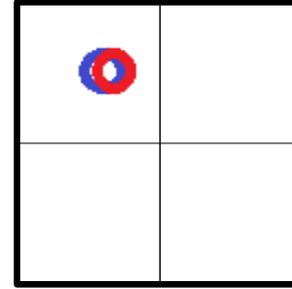
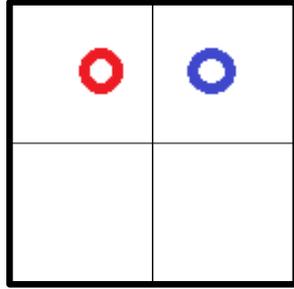
Profil



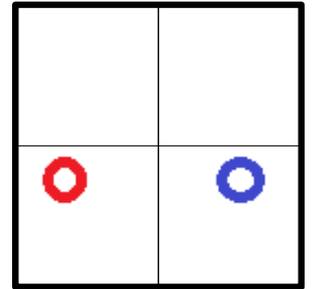
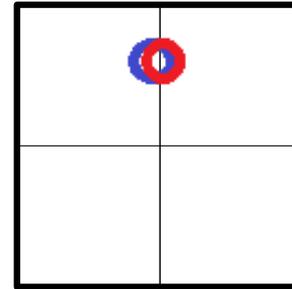
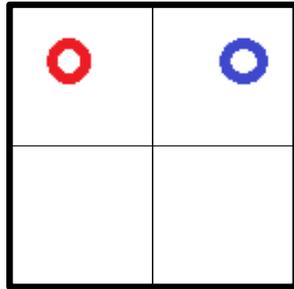
Dessus



T=1



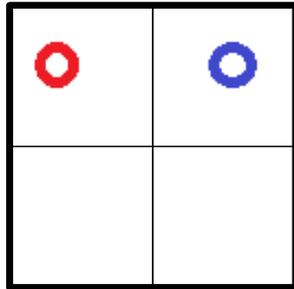
T=2



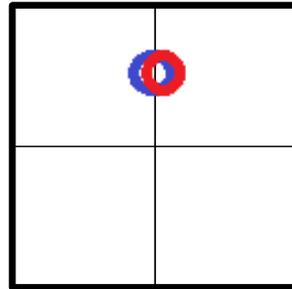
Action: Zoom arrière

Nom du geste: Fermeture ascenseur

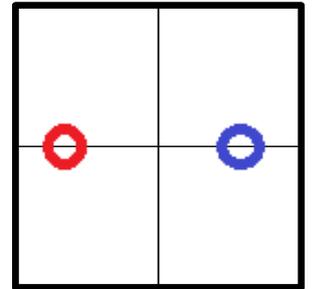
T=0



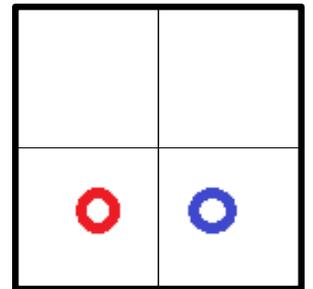
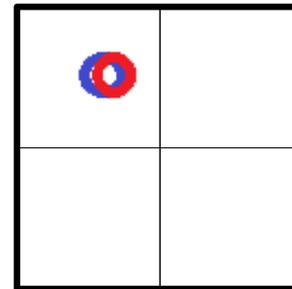
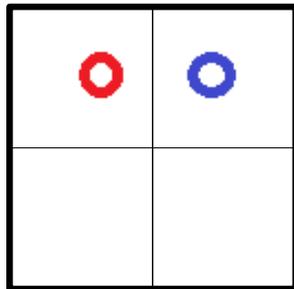
Profil



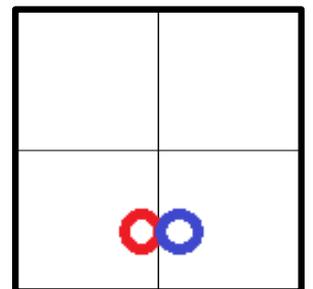
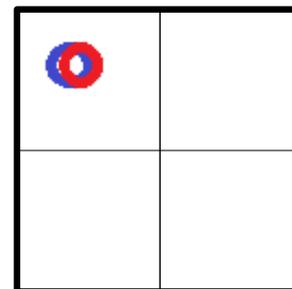
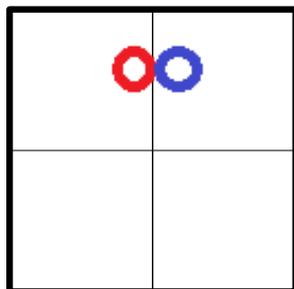
Dessus



T=1



T=2



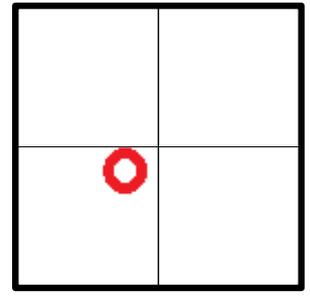
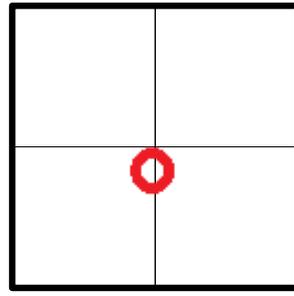
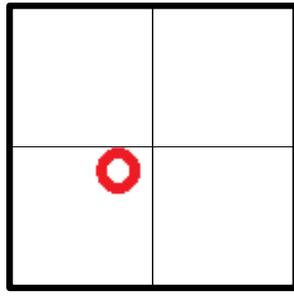
Action: Aller dans Sreet View

Diagonale main droite

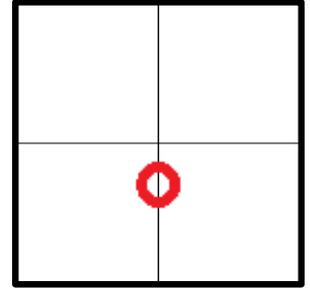
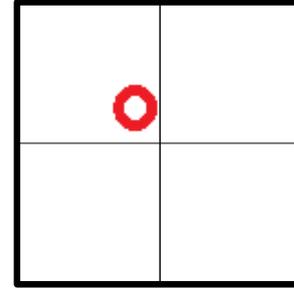
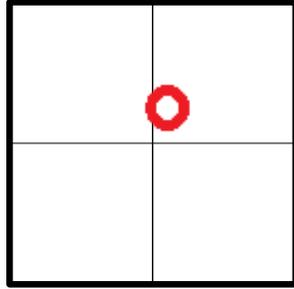
montante vers gauche

Nom du geste:

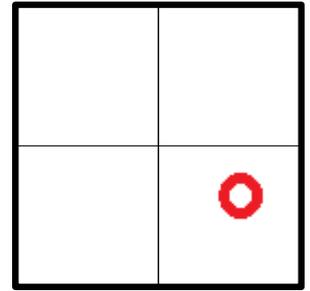
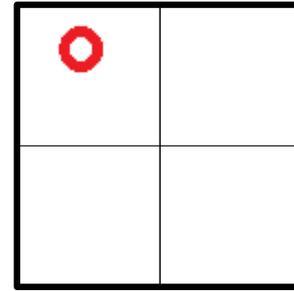
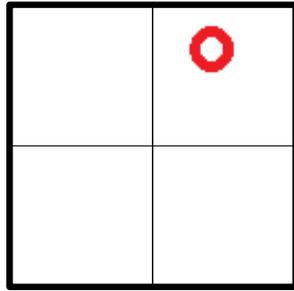
T=0



T=1



T=2



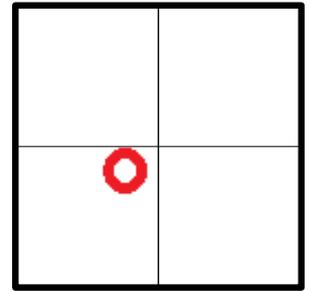
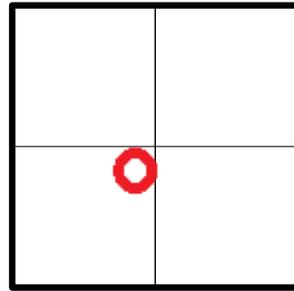
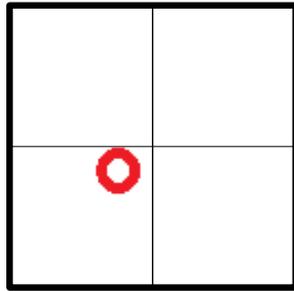
Action: Aller dans Google Map

Diagonale main droite

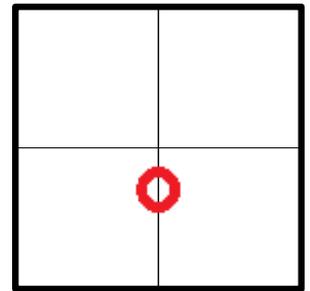
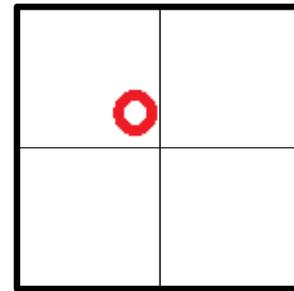
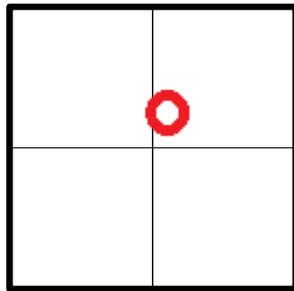
montante vers gauche

Nom du geste:

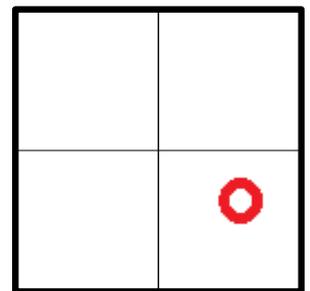
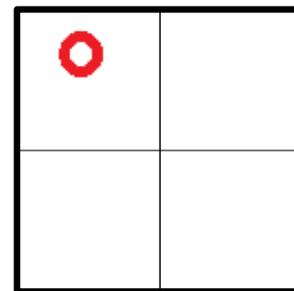
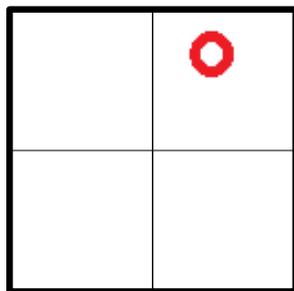
T=0



T=1



T=2

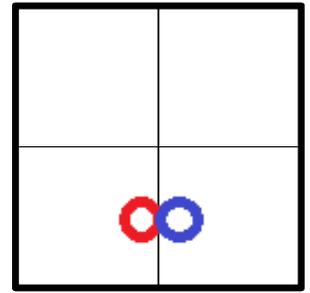
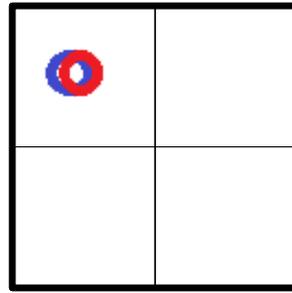
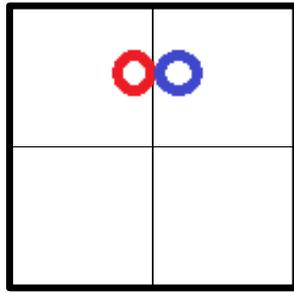


Face

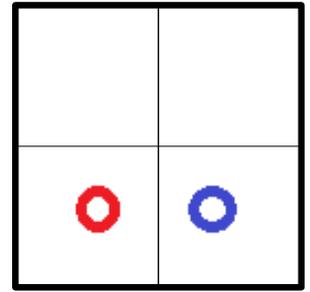
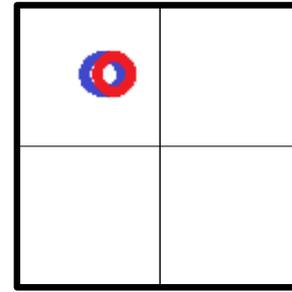
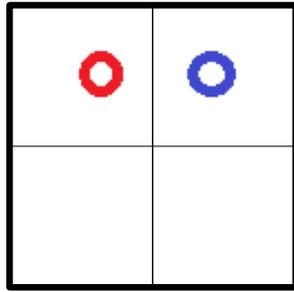
Profil

Dessus

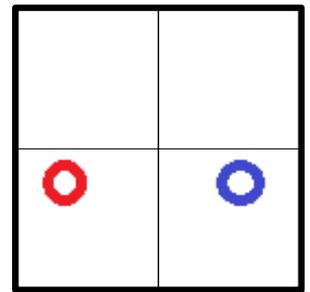
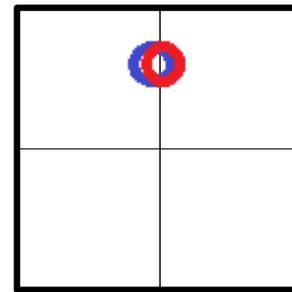
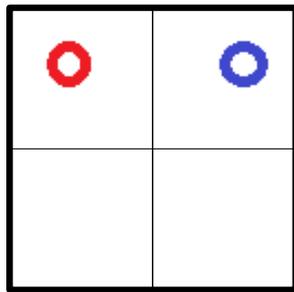
T=0



T=1



T=2



Action: Avancer

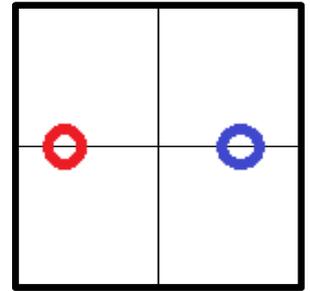
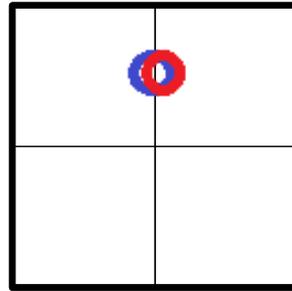
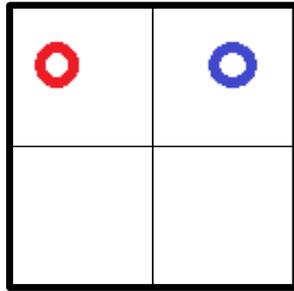
Nom du geste: Ouverture ascenseur

Face

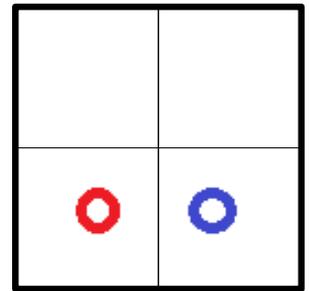
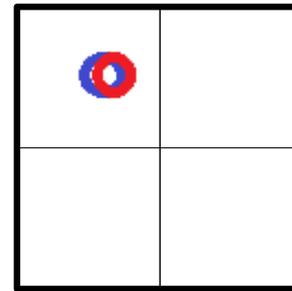
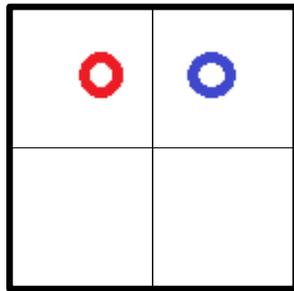
Profil

Dessus

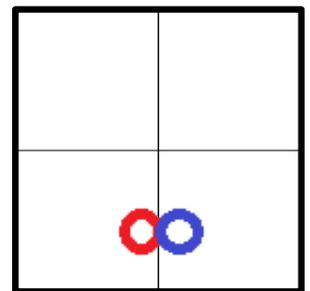
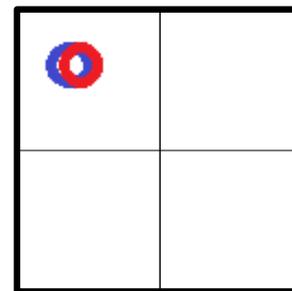
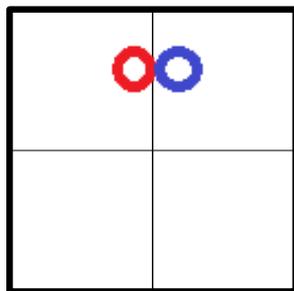
T=0



T=1



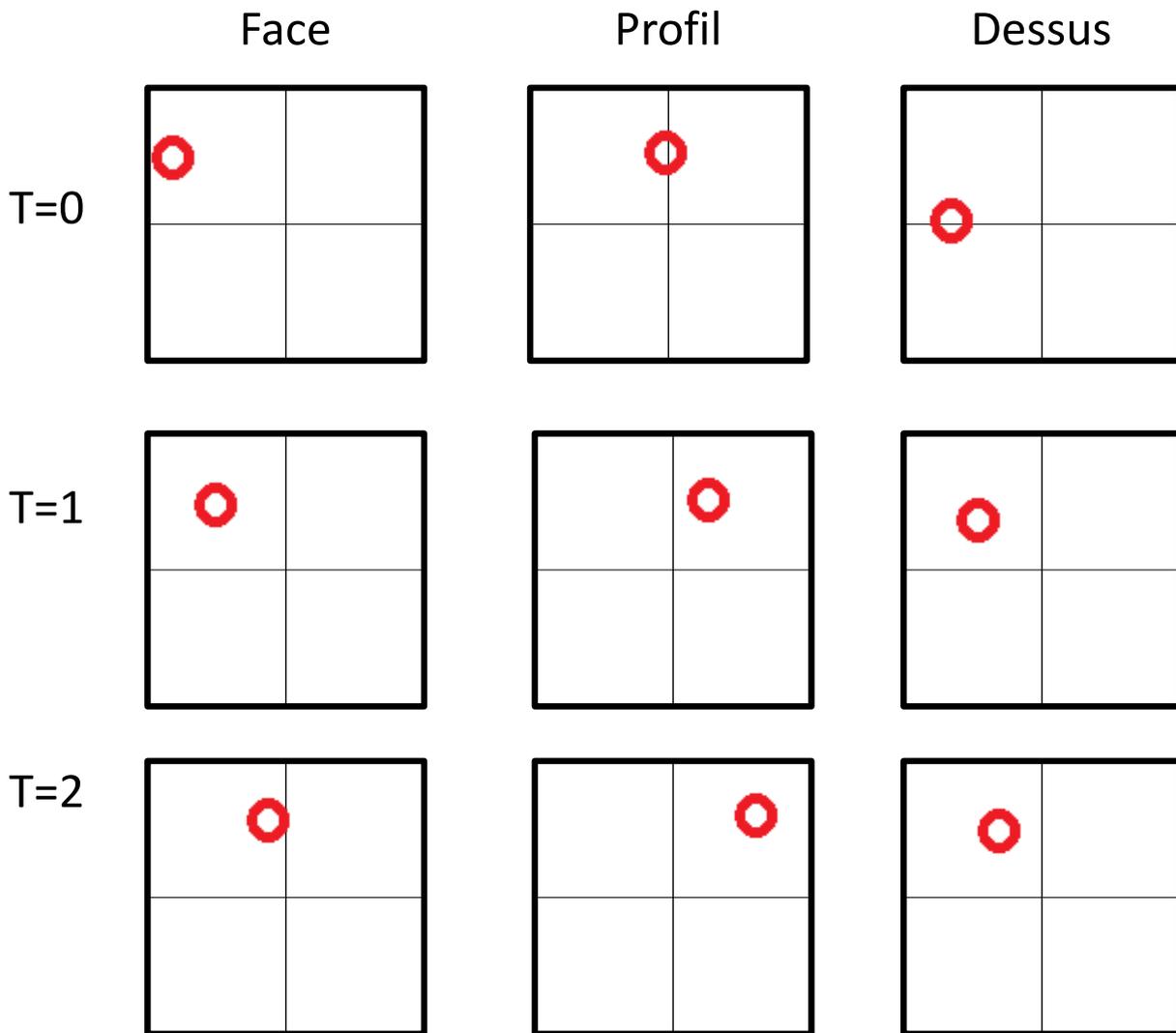
T=2



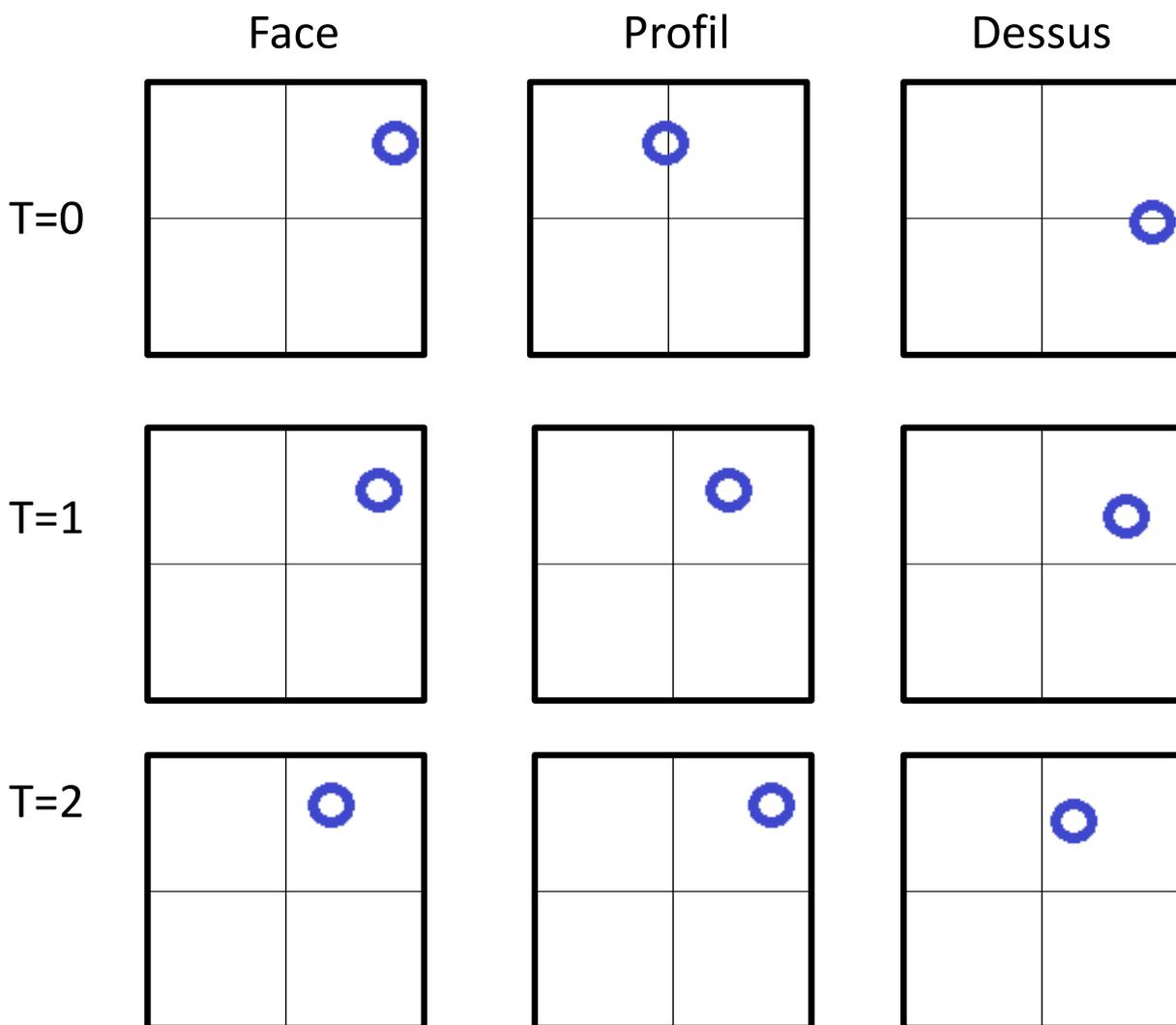
Action: Reculer

Nom du geste: Fermeture ascenseur

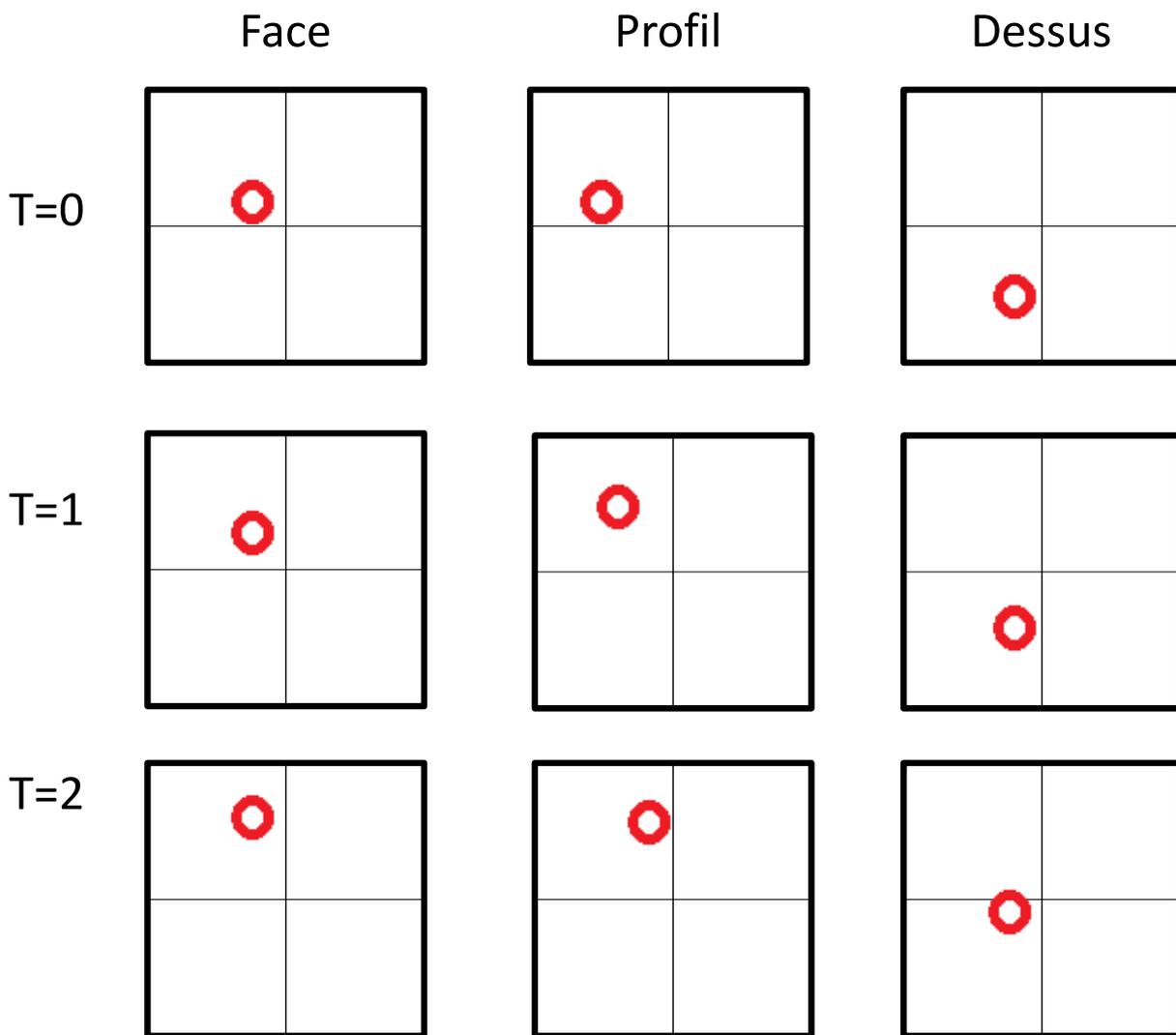
Action: Aller à droite
 Nage main droite sans passer devant



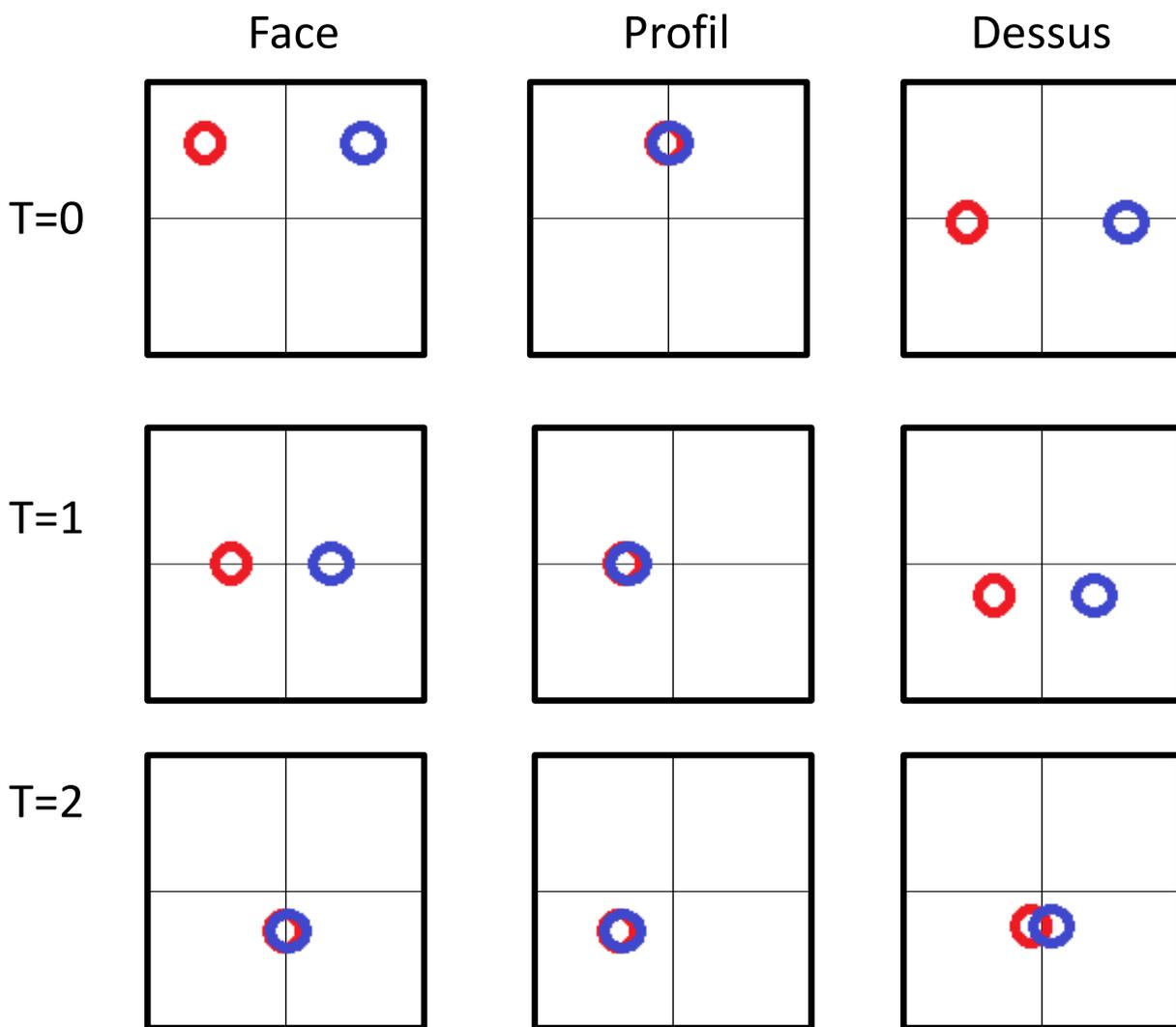
Action: Aller à gauche
 Nage main gauche sans passer devant



Action: Regarder en haut
 Nom du geste: Monte haltère main droite



Action: Regarder en bas
 Rejoint les mains vers le bas
 bras tendu
 Nom du geste:



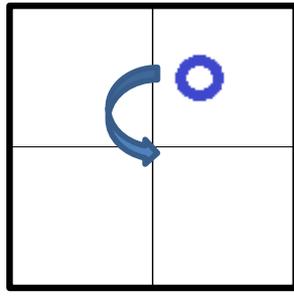
Gestes FLOP

Action: Zoom avant

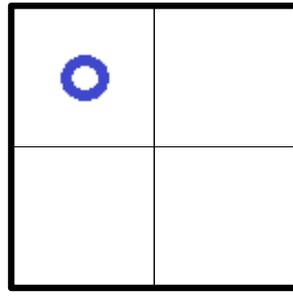
Cercle main gauche bras tendu
sens horaire

Nom du geste:

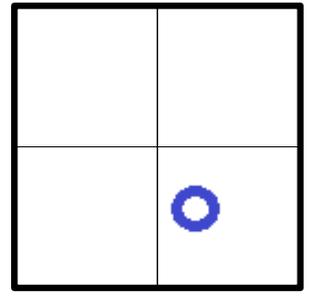
T=0



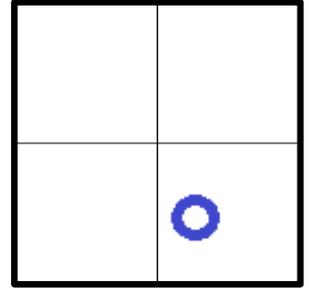
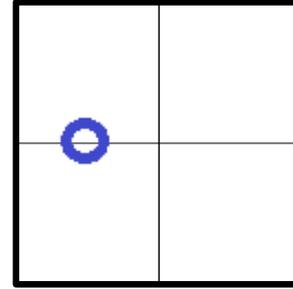
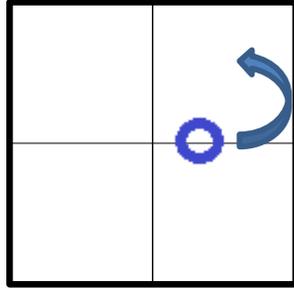
Profil



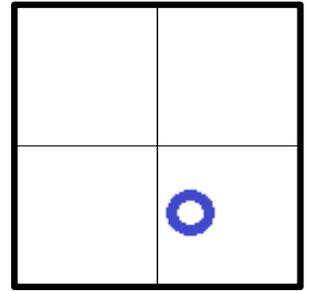
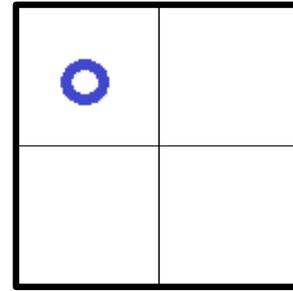
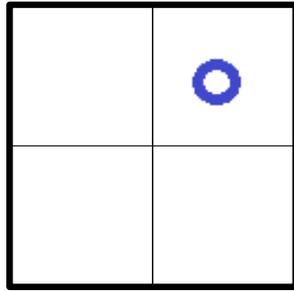
Dessus



T=1



T=2

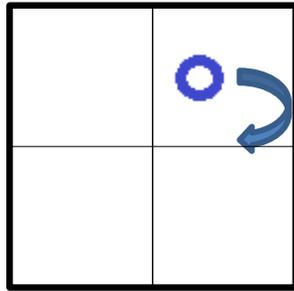


Action: Zoom arrière

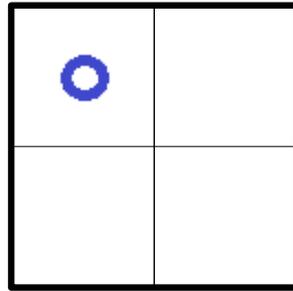
cercle main gauche bras
tendu anti horaire

Nom du geste:

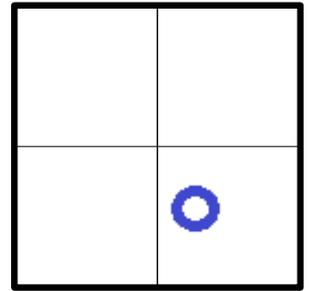
T=0



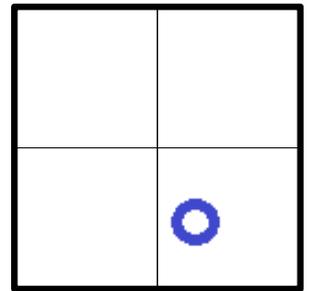
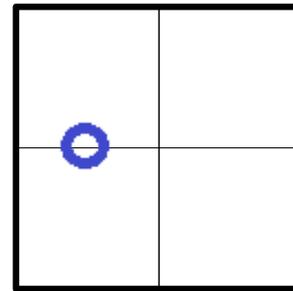
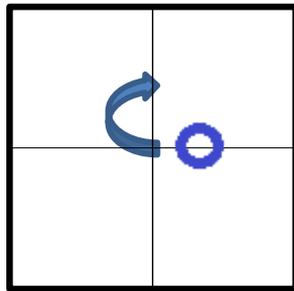
Profil



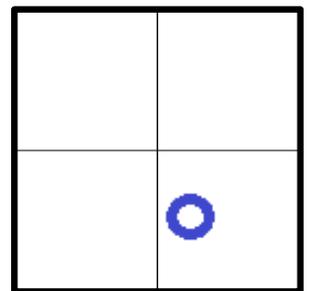
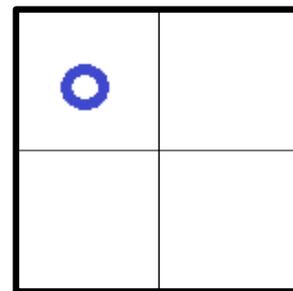
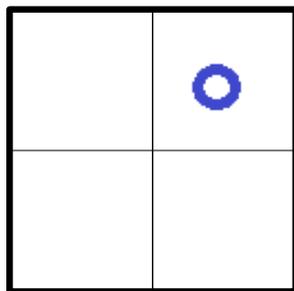
Dessus



T=1



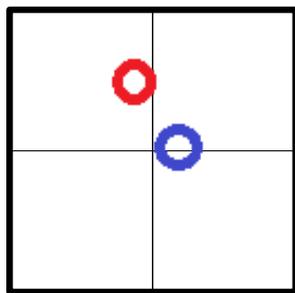
T=2



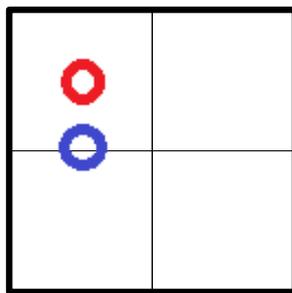
Action: Aller dans Sreet View

Nom du geste: Moulinette deux mains
sens horaire

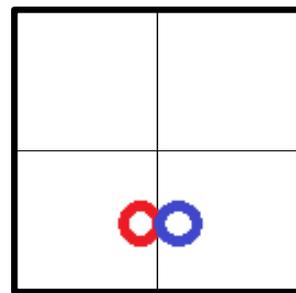
T=0



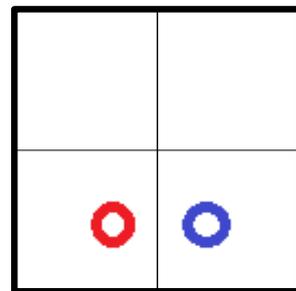
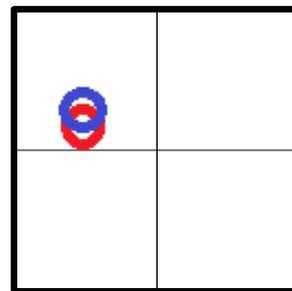
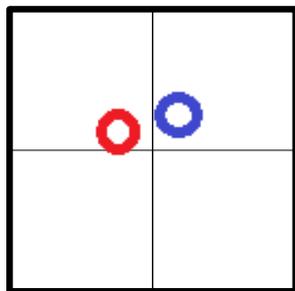
Profil



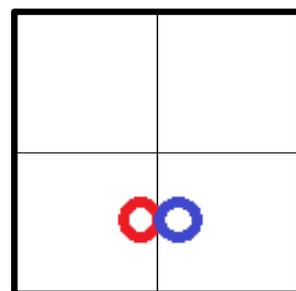
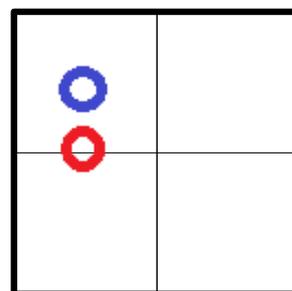
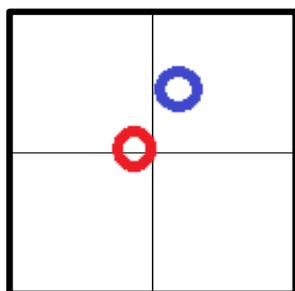
Dessus



T=1



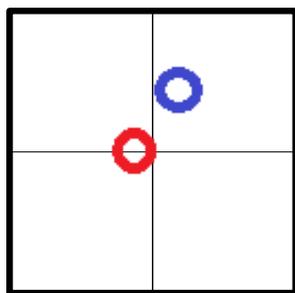
T=2



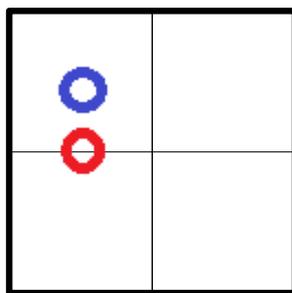
Action: Aller dans Google Map

Nom du geste: Moulinette deux mains
sens anti horaire

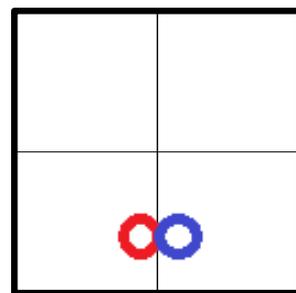
T=0



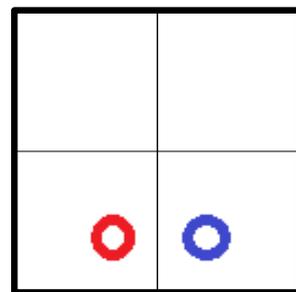
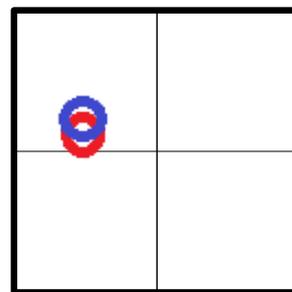
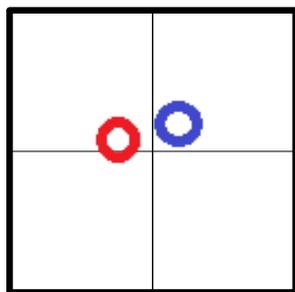
Profil



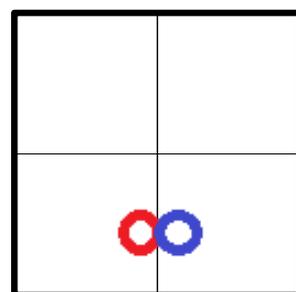
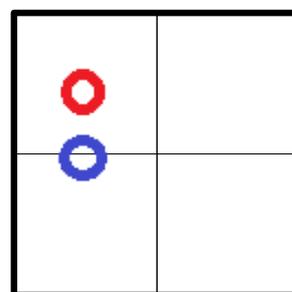
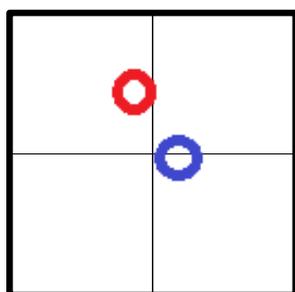
Dessus



T=1



T=2



Gestes prototypiques

Face

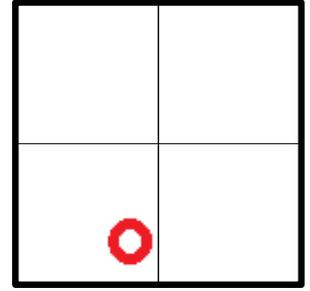
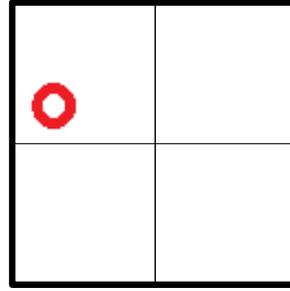
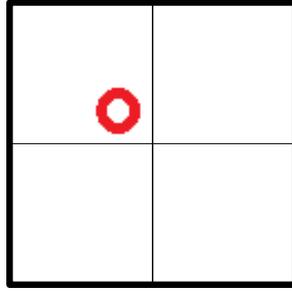
Profil

Dessus

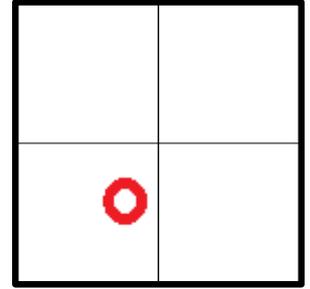
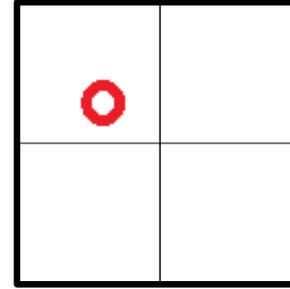
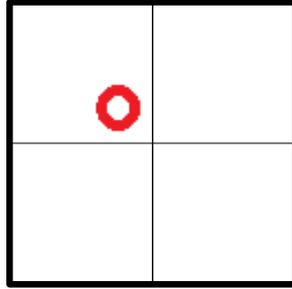
Action: Avancer

Nom du geste: Tirer main droite

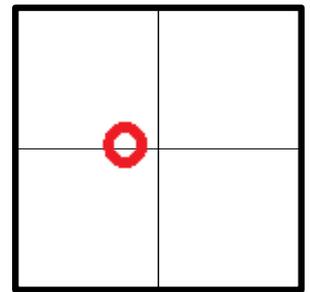
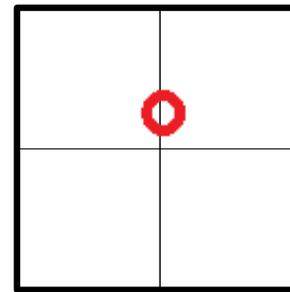
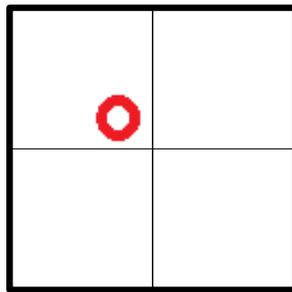
T=0



T=1



T=2



Face

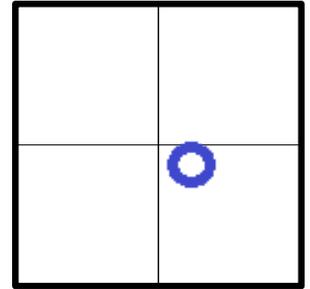
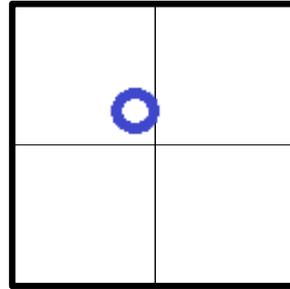
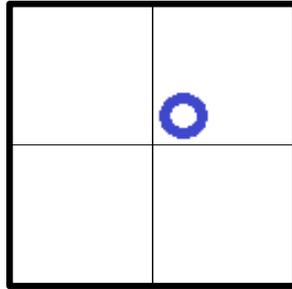
Profil

Dessus

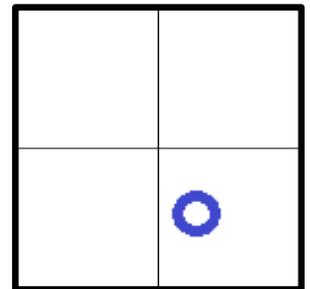
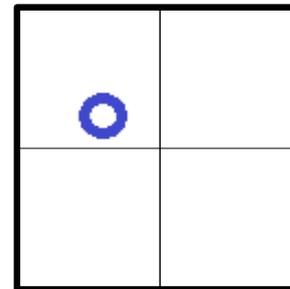
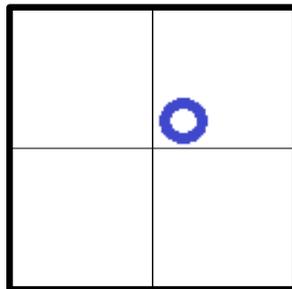
Action: Reculer

Nom du geste: Pousser main gauche

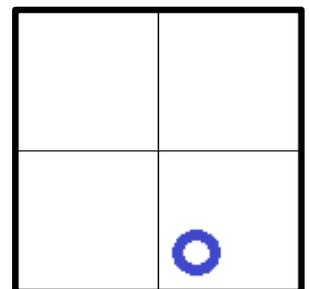
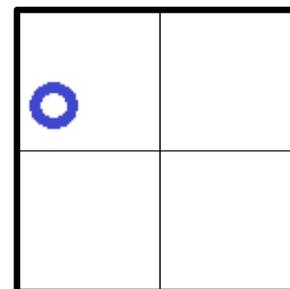
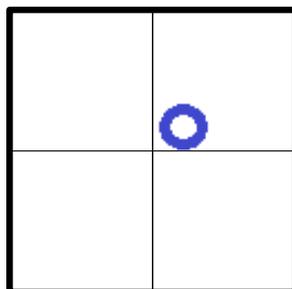
T=0



T=1



T=2



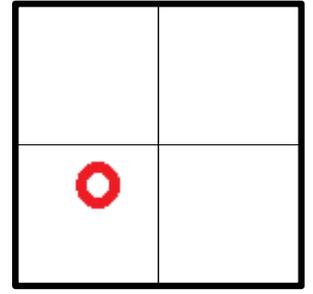
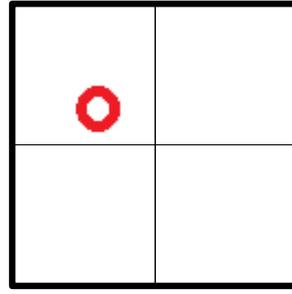
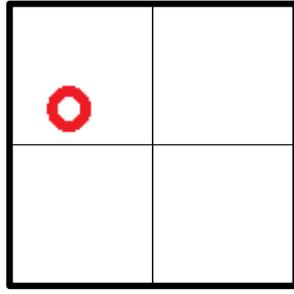
Gestes prototypiques

Face

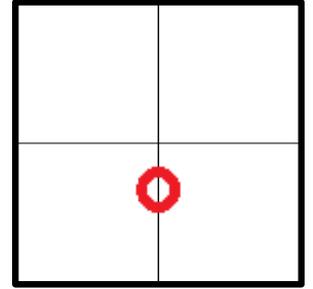
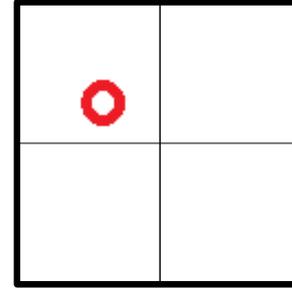
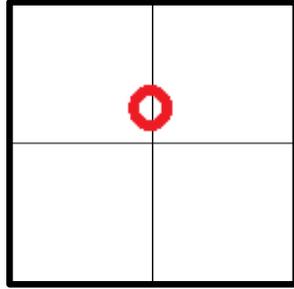
Profil

Dessus

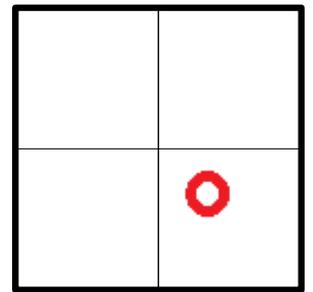
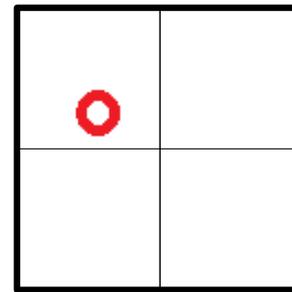
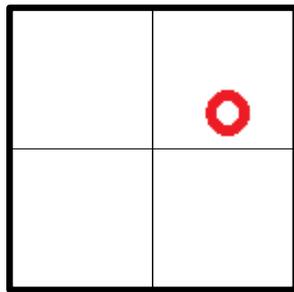
T=0



T=1



T=2



Action: Aller à droite

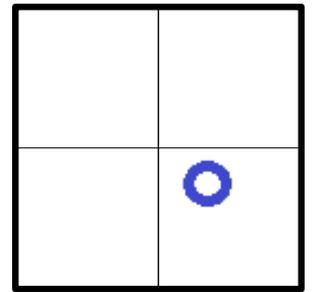
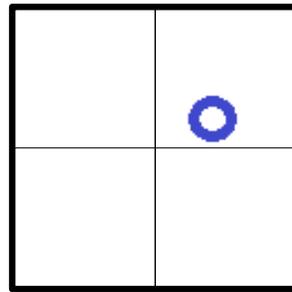
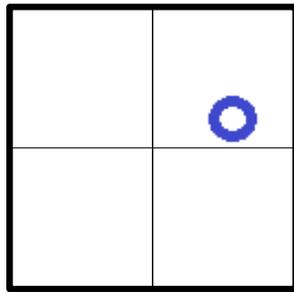
Nom du geste: Swipe left

Face

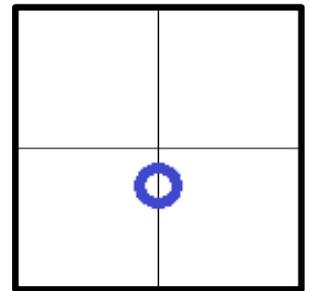
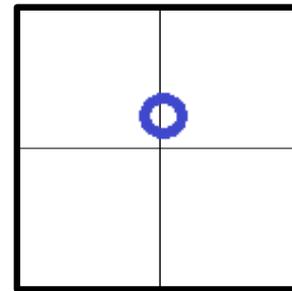
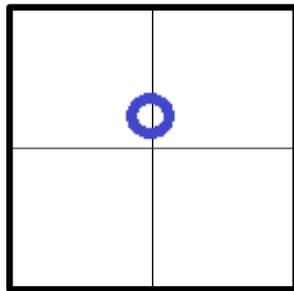
Profil

Dessus

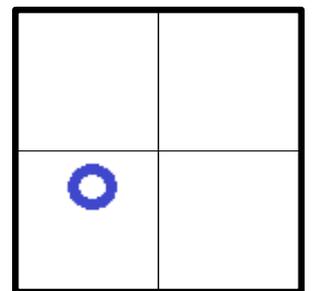
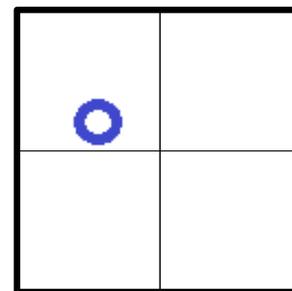
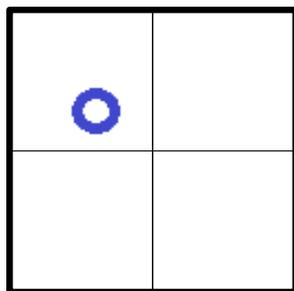
T=0



T=1



T=2



Action: Aller à gauche

Nom du geste: Swipe right

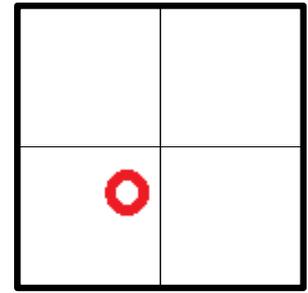
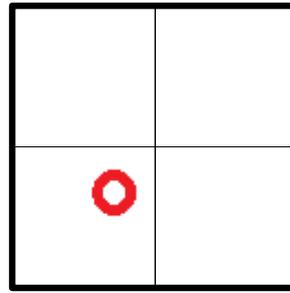
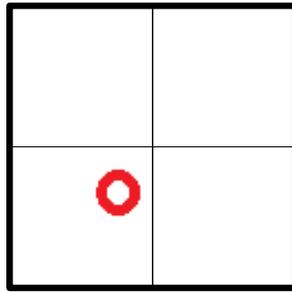
Gestes prototypiques

Face

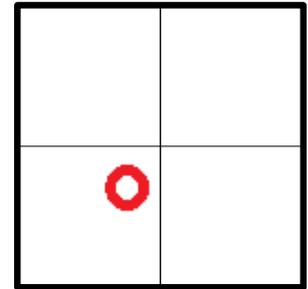
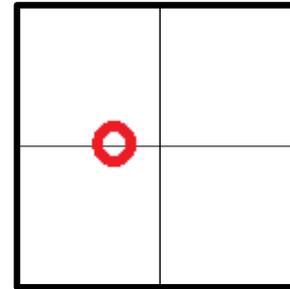
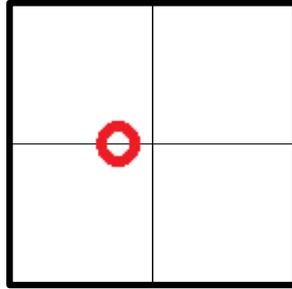
Profil

Dessus

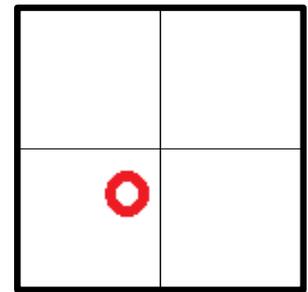
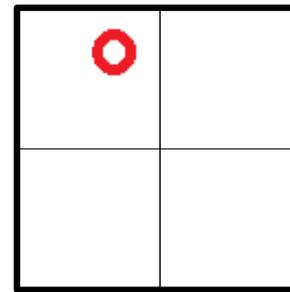
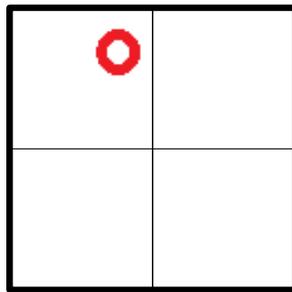
T=0



T=1



T=2



Action: Regarder en haut

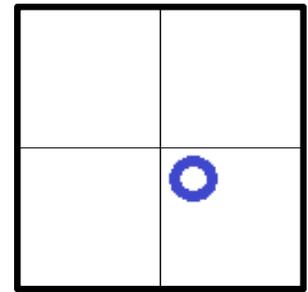
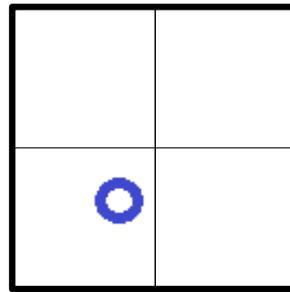
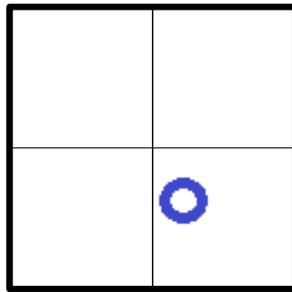
Nom du geste: Lorem ispum

Face

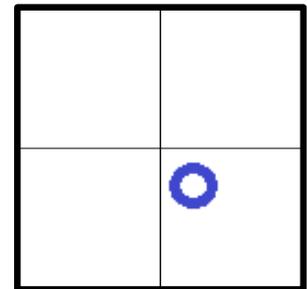
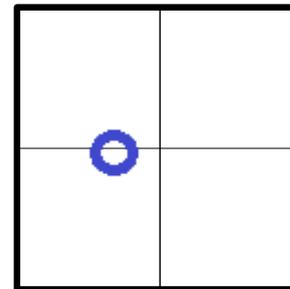
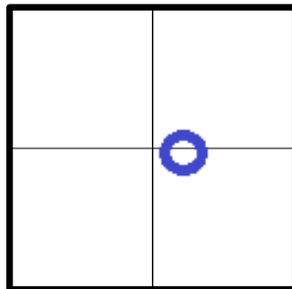
Profil

Dessus

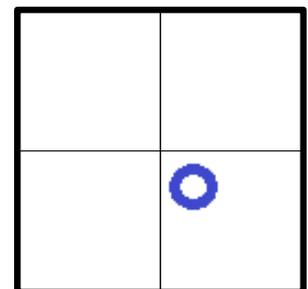
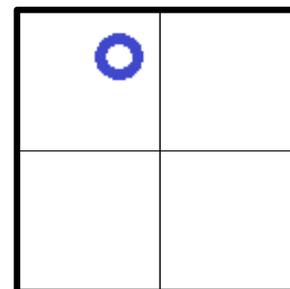
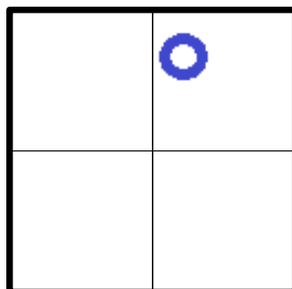
T=0



T=1



T=2



Action: Regarder en bas

Nom du geste: Lorem ispum

Gestes prototypiques

Face

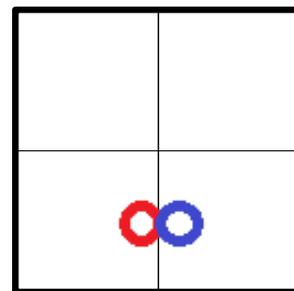
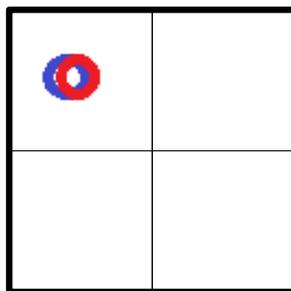
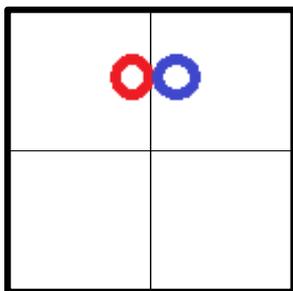
Profil

Dessus

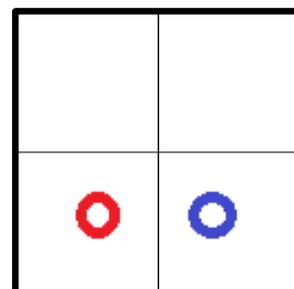
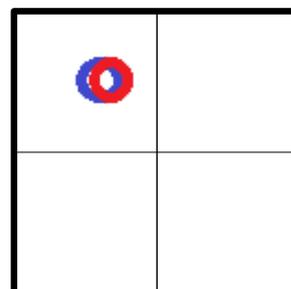
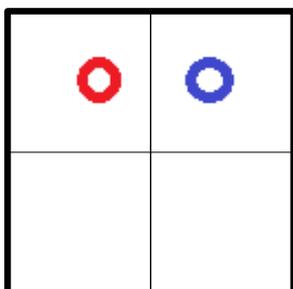
Action: Zoom avant

Nom du geste: Ouverture ascenseur

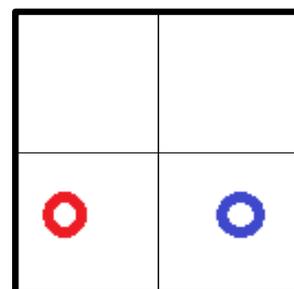
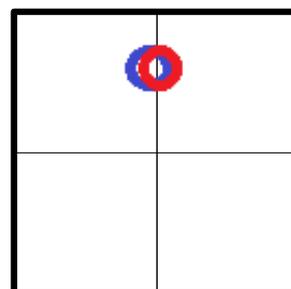
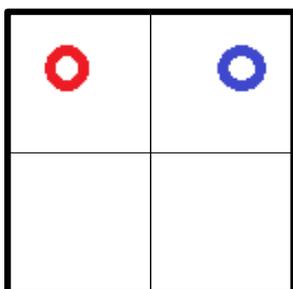
T=0



T=1



T=2



Face

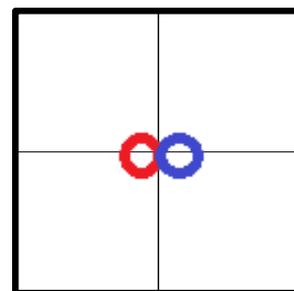
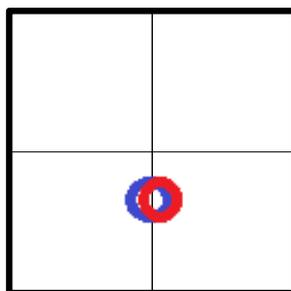
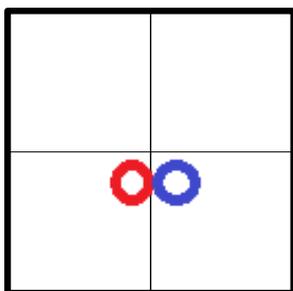
Profil

Dessus

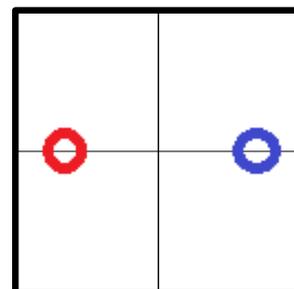
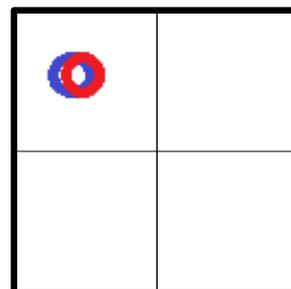
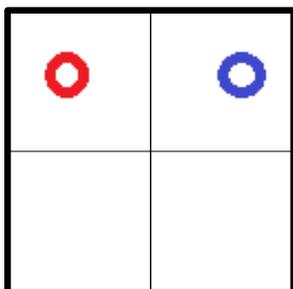
Action: Zoom arrière

Nom du geste: Deux mains devant corps + fermeture ascenseur

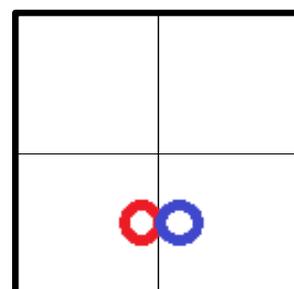
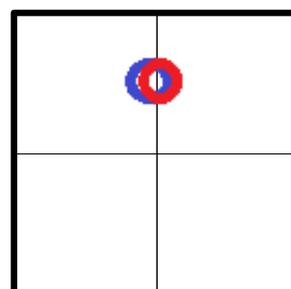
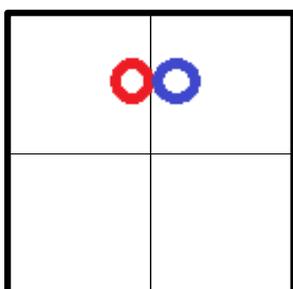
T=0



T=1



T=2



Gestes prototypiques

Face

Profil

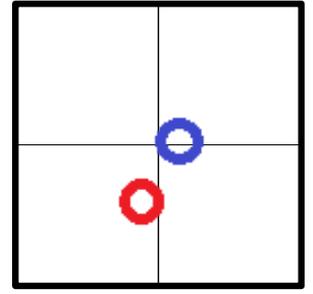
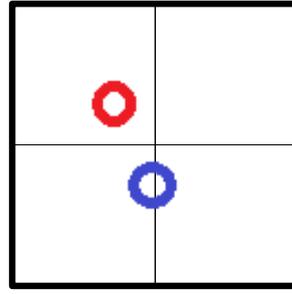
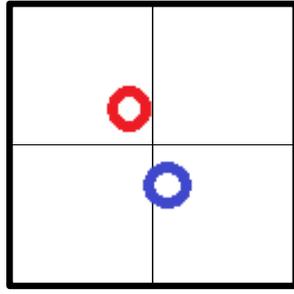
Dessus

Action: Aller dans Street View

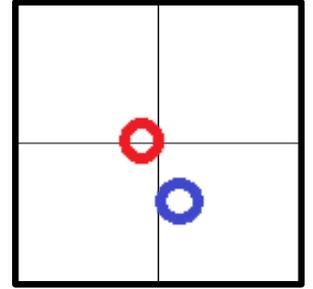
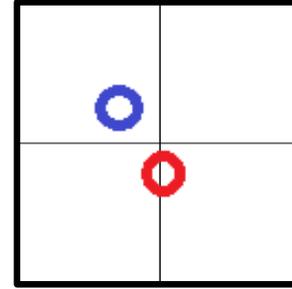
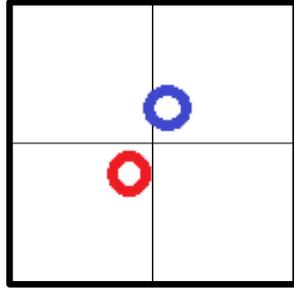
Mains minent mouvement

Nom du geste: course

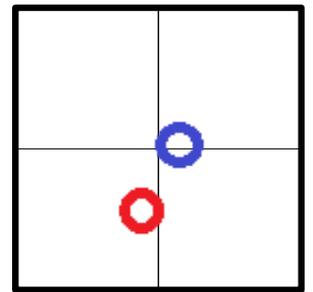
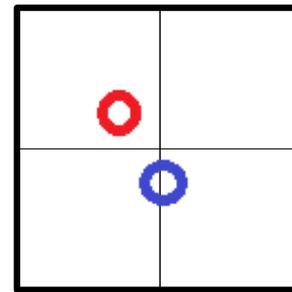
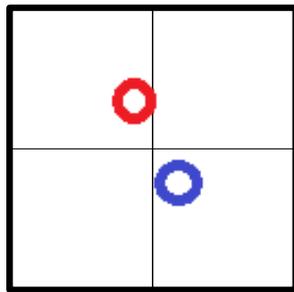
T=0



T=1



T=2



Face

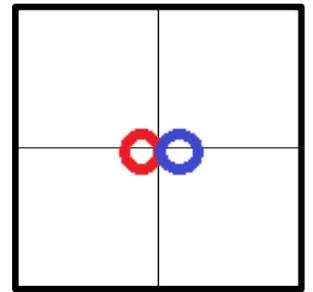
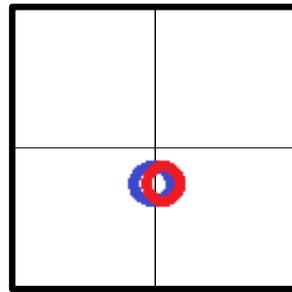
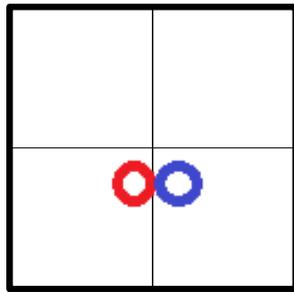
Profil

Dessus

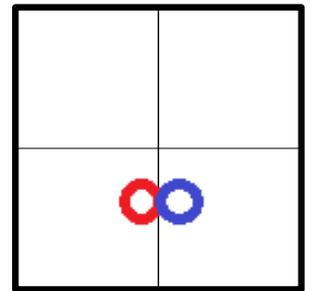
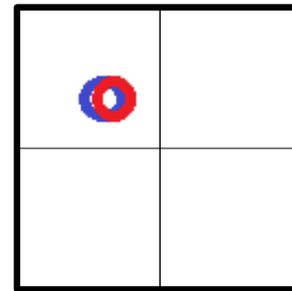
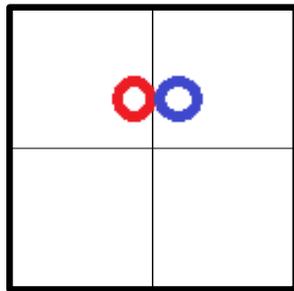
Action: Aller dans Google Map

Nom du geste: Dépliage carte routière

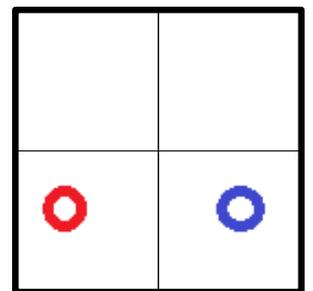
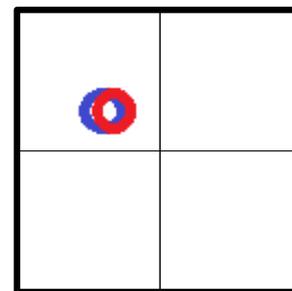
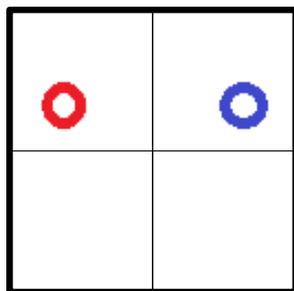
T=0



T=1



T=2



Les gestes des vingt-cinq sujets et leurs réponses aux questions sont disponibles sur Dropbox à l'adresse suivante :

<https://www.dropbox.com/sh/nxizveq5s7s1r4t/ZcgHC0Y4UM>

Les questionnaires effectués à Renaissance sont également à cette même adresse.