



Compte rendu

Projet IHM Multimodal

**Création d'un moteur multimodal
de type "Put there there"**

FOUQUART Baptiste
PRZYJEMSKI Kosma
HERNANDEZ Zamayra

Modélisation du moteur multimodal

Intro

Le “*Projet IHM Multimodale*” a pour le but de créer une interface de type “*Put that there*” qui est opérée par la fusion de la reconnaissance vocale, de la reconnaissance de gestes et du pointage de souris. Originellement, le projet “*Put that there*” a été conduit par MIT en 1982. Avec l’avancement de la technologie, la reprise de ce projet permet de faire face à l’implémentation du plusieurs modes d’interaction entre l’homme et la machine en utilisant des utiles facilement disponibles qui ne demandent pas beaucoup de temps à développer. Le projet est donc un prototype fonctionnel du moteur multimodal permettant la gestion de formes géométriques avec des actions supplémentaires ajoutées.

Chronogramme

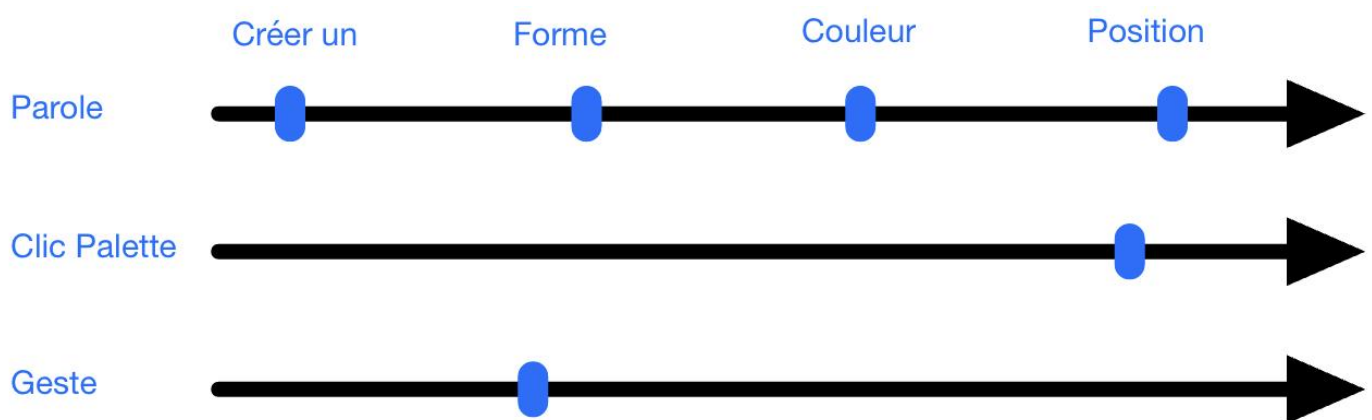


Fig. 1 - Chronogramme du moteur multimodal.

On voit sur le chronogramme qu’on demande au moteur de créer une forme grâce à une commande vocale. Ensuite, on va devoir choisir le type de forme. Pour cela, le moteur analyse d’abord si l’information a été donnée via un geste effectué avec la souris, et reconnu grâce au logiciel *ICAR*, que nous avons entraîné, ou alors, on choisit avec une nouvelle commande vocale. La reconnaissance de la parole est effectuée par le logiciel *sra5*. La couleur est également choisie via reconnaissance vocale. Finalement, la position est soit déterminée par un clic de la souris aux coordonnées souhaitées, ou en donnant une position relative à un autre objet vocalement.

Conception

La reconnaissance de gestes a été faite grâce au logiciel *ICAR* pour. Ce logiciel permet de reconnaître des formes à partir d’une base de données d’apprentissage. Nous avons donc créé des bases de données pour les formes que l’on veut reconnaître, à savoir triangle, rectangle, losange et cercle. Nous avons rendu l’apprentissage plus robuste en insérant des modèles qui n’étaient pas parfait. Cela veut dire qu’on a noté comme formes correctes des formes qui étaient relativement mal dessinées. Cette idée a rendu la reconnaissance très robuste aux formes dessinées rapidement (imparfaites), cependant nous avons rencontré un problème puisque les formes dessinées avec beaucoup d’application n’étaient souvent plus reconnues correctement. Nous avons alors supprimé des modèles imparfaits de la base de données pour rétablir un équilibre dans la reconnaissance des formes.

La reconnaissance vocale a été effectuée avec le logiciel *sra5*. Nous avons modifié la grammaire de base *grammar.grxml* pour ajouter de l’option “changer couleur”. Nous avons également ajouté certains

éléments dans la grammaire afin de rendre le moteur multimodal plus “user friendly”, notamment en donnant à l'utilisateur plus de liberté d'expression à l'orale. La grammaire a été testée au début avec les mots reconnus, et pour l'implémentation en processing on a décidé d'envoyer les *concepts* pour une meilleure utilisation des variables.

La communication entre la reconnaissance des gestes et la reconnaissance vocale est faite à partir des buses *ivy*, qui reçoivent les concepts de *icar* et *sra5*, et les renvoient sur l'application processing.

Machine à état

Le fonctionnement du moteur multimodal est modélisé par la machine à état suivante :

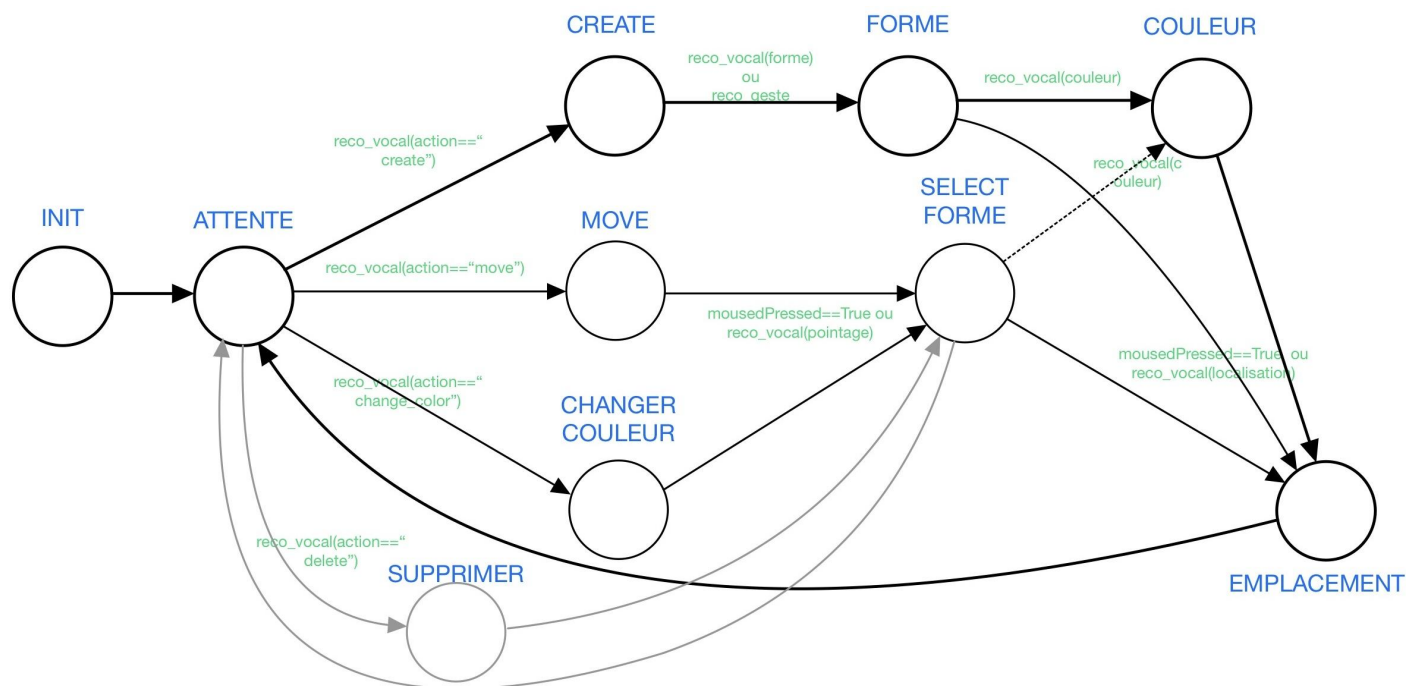


Fig. 2 - Machine à état du moteur multimodal.

Voici l'explication du fonctionnement “pas par pas” du fonctionnement du système de point de vu d'utilisateur :

1. L'utilisateur dit le nom d'une action parmi : “créer (+un)”, “déplacer”, “changer couleur” ou “supprimer”.
2. Pour l'action “créer (+un)” :
 1. L'utilisateur dessine un geste sur l'interface de Icarivy ou il dit le nom d'une forme parmi : “rectangle”, “triangle”, “cercle” ou “losange”.
 2. Optionnellement : pour sélectionner un couleur, l'utilisateur dit le nom d'un couleur parmi : “blue”, “rouge”, “orange”, “vert”, “jaune”, “violet” ou “noire”.
 3. L'utilisateur spécifie la destination en cliquant sur l'endroit d'intérêt sur l'interface graphique de la palette ou en disant “ici”, “là” ou “à cet endroit” et pointant sur l'endroit d'intérêt. L'objet de la forme choisie, optionnellement du couleur choisi, devrait être créé à l'endroit spécifié par l'utilisateur.
3. Pour l'action “déplacer” :
 1. L'utilisateur spécifie la forme en cliquant sur l'objet à déplacer sur l'interface graphique de la palette ou en disant “ici”, “là” ou “à cet endroit” et pointant sur l'objet désiré.
 2. L'utilisateur spécifie la destination en cliquant sur l'endroit d'intérêt sur l'interface graphique

de la palette ou en disant "ici", "là" ou "à cet endroit" et pointant sur l'endroit d'intérêt.

L'objet choisi devrait être déplacé à une nouvelle destination choisie par l'utilisateur.

4. Pour l'action "changer couleur" :

1. L'utilisateur spécifie la forme en et pointant sur l'objet désiré.

2. Optionnellement : pour sélectionner un couleur, l'utilisateur dit le nom d'un couleur parmi : "blue", "rouge", "orange", "vert", "jaune", "violet" ou "noire".

L'objet choisi devrait changer de couleur, soit au couleur précise par l'utilisateur, soit a un couleur random.

5. Pour l'action "supprimer" :

1. L'utilisateur doit spécifier la forme a supprimer en pointant sur l'objet a enlever avant qu'il prononce la commande. L'objet choisi par le pointage devrait être enlevé de la palette.

Pour l'action "créer (+un)", l'utilisateur peut choisir entre prononcer "créer" et "créer un".

Pour l'action "déplacer", l'utilisateur peut aussi dire "bouger ça ici".

Pour l'action "changer couleur" on peut également dire "changer couleur de ca".

Pour l'action "supprimer" on peut utiliser les expressions suivantes : "supprimer", "supprimer cette forme" ou "enlever".

Exemples d'utilisation

Veuillez regarder les vidéos de démonstration d'utilisation du moteur multimodal sous le lien

<https://clipchamp.com/watch/rhX3aEli1Gd>

Une instruction d'utilisation du moteur multimodal est disponible dans le ReadMe du projet qui se trouve dans répertoire contenant le code.