

Multimodal Interface Project

HomeControl

Loïc Baechler - Jonathan Sifringer - Pierre Tinguely

4/26/2013

Table of Contents

| | | |
|----|--|---|
| 1 | Introduction | 2 |
| 2 | But | 2 |
| 3 | Description de l'application | 2 |
| 4 | Maquette de l'application | 3 |
| 5 | Modalité | 4 |
| 6 | Matériels | 4 |
| 7 | Software | 4 |
| 8 | Fusion et Fission..... | 5 |
| | Fusion | 5 |
| | Fission..... | 5 |
| 9 | Points de vue CASE/CARE..... | 5 |
| | CASE | 5 |
| | CARE | 5 |
| 10 | Tableau d'interaction | 6 |
| 11 | Problèmes rencontrés et limitations..... | 7 |
| 12 | Conclusion | 7 |

1 Introduction

Dans le cadre du cours Multimodal Interface, on est amené par groupe de 3 à réaliser une application utilisant différentes modalités. On a pensé rendre une pièce d'un appartement intelligente et réagissant à l'aide de la voix et de la présence à l'intérieur de la pièce.

2 But

Représenter virtuellement l'emplacement d'une personne dans une pièce. Différents objets se trouvant dans la pièce pourront alors être commandé de deux façons possibles. A l'aide de la voix uniquement, ou alors en fonction de la localisation de la personne dans la pièce et la voix.

3 Description de l'application

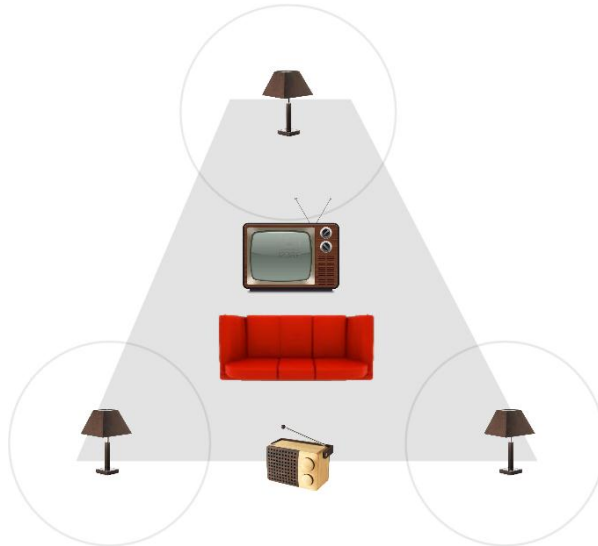
Une personne se trouvera dans une pièce filmée par la Kinect. Son emplacement sera représenté par un cercle à l'écran.

Des objets seront positionnés dans la pièce par exemple une lampe, une radio, et une télévision. Il sera possible à l'aide de la Kinect, d'allumer et d'éteindre les différents périphériques à condition de se trouver dans la pièce ou à un emplacement précis (par exemple sur le canapé pour allumer la télévision). L'utilisateur devra aussi utiliser la voix pour interagir avec les différents périphériques dans les salles. Pour la télévision, l'utilisateur devra simplement s'asseoir sur le canapé pour qu'elle s'allume et se relever pour qu'elle s'éteigne.

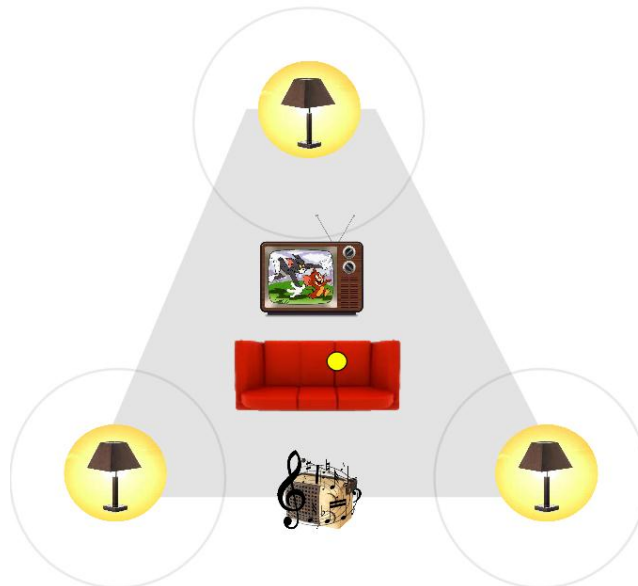
4 Maquette de l'application

Nous avons créé une interface graphique 2D qui représente une salle qui est entièrement couverte par la Kinect. La partie en gris représente approximativement le champ de vision de la Kinect.

Cette image représente la salle et tous ses périphériques éteints:



Cette image représente la salle et tous ses périphériques allumés:



5 Modalité

On pourra interagir avec l'application de plusieurs manières:

- Par la présence et l'emplacement de la personne dans la pièce et à l'aide de commandes vocales
- Par des commandes vocales uniquement
- Par des postures (assis ou pas)
- Emplacement de l'utilisateur

Pour pouvoir capturer ces différentes modalités nous avons choisi d'utiliser la "Kinect for XBOX 360" car elle nous permet d'une part de capturer les mouvements et les postures mais l'API nous donne également la possibilité d'analyser les mots et les phrases énoncées à haute voix en français.



La Kinect contient un microphone et une caméra pour détecter notre emplacement et la position de notre corps.

Nos modalités sont donc la voix, la posture et l'emplacement géographique

6 Matériels

- Kinect for Xbox 360
- Ordinateur

7 Software

- Windows 7
- Microsoft Visual Studio 2010
- ASP.NET

- Kinect for windows Speech Recognition Language Pack (fr-FR)
- Kinect for Windows SDK v1.6

8 Fusion et Fission

Fusion

Le Framework du capteur Kinect nous transmet les informations de localisation, de posture ainsi que le traitement de la voix. Par la suite notre application procède à la computation pour générer les sortie, voir fission.

Fission

Notre application est très discrète, nous avons gardé le feedback de l'application au minimum pour ne pas gêner l'expérience de l'utilisateur. Lorsque l'application détecte une entrée valide, elle va allumer le bon périphérique et c'est tout ce dont l'utilisateur a besoin.

9 Points de vue CASE/CARE

CASE

Selon le type de communication multimodale CASE, notre projet est "alternative" et "exclusive". Lorsque nous allons dans l'espace d'un périphérique à allumer et que nous disons "Allume", cela va allumer le périphérique en question. Cette propriété est dite alternative car nous avons la modalité de la localisation, ainsi que celle de la reconnaissance vocale. Il est de même pour la télévision qui va s'allumer lorsque nous nous asseyons sur le canapé.

Pour le type "exclusive", nous pouvons allumer les périphériques de la salle en disant "allume X de la salle Y". Cela aura pour effet d'activer le périphérique en question depuis n'importe quelle localité.

CARE

Selon le type de communication multimodale CARE, notre projet est "Equivalence" et "Complementarity". Complémentaire car nous pouvons allumer ou éteindre un périphérique en étant dans son espace ou nous pouvons l'allumer ou éteindre depuis n'importe où en épelant son nom et la directive.

Il utilise aussi le modèle équivalent car nous pouvons allumer la télévision soit en indiquant la directive avec la voix (Vocal), soit en nous asseyant sur le canapé (posture).

10 Tableau d'interaction

| <u>Type de Périphériques</u> | <u>Position géographique</u> | <u>Action vocale</u> | <u>Effet</u> |
|------------------------------|------------------------------|----------------------|--|
| Tous | Pas dans la pièce | - | Tous les périphériques vont s'éteindre |
| Tous | N'importe où dans la pièce | Allume tout | Tous les périphériques vont s'allumer |
| Télévision | N'importe où dans la pièce | Éteins la télévision | Éteins la télévision |
| Télévision | N'importe où dans la pièce | Allume la télévision | Allume la télévision |
| Télévision | S'asseoir sur le canapé | - | Allume la télévision |
| Télévision | Se lever du canapé | - | Éteins la télévision |
| Lampe de X | Proche de la lampe | Éteins | Éteins la lampe de X |
| Lampe de X | Proche de la lampe | Allume | Allume la lampe de X |
| Radio | N'importe où dans la pièce | Éteins la radio | Éteins la radio |
| Radio | N'importe où dans la pièce | Allume la radio | Allume la radio |
| Radio | Proche de la radio | Allume | Allume la radio |
| Radio | Proche de la radio | Éteins | Éteins la radio |
| Lampe de X | N'importe où dans la pièce | Allume la lampe de X | Allume la lampe de X |
| Lampe de X | N'importe où dans la pièce | Éteins la lampe de X | Éteins la lampe X |

11 Problèmes rencontrés et limitations

Notre application peut avoir des difficultés de reconnaissance vocal lorsqu'il y'a trop de bruits alentour (par exemple si la musique est trop forte). Certaines fois, la reconnaissance vocale comprendra quelque chose que nous n'avons pas dit et fera une action délibérément.

La Kinect a un champ de vision assez limité et empêcherait le bon fonctionnement de notre application dans une grande salle.

12 Conclusion

Durant le projet, il a été intéressant de travailler avec les différentes modalités que nous avons vues en théorie. Notre projet utilisant la voix et le corps est un prototype d'une application qui pourrait être utilisé dans de vraies salles. Il faudrait par contre que la Kinect ait une plus grande vision et que la librairie utilisée ait un meilleur moteur de reconnaissance vocal. Dans notre application, nous représentons les différents périphériques virtuellement mais on pourrait utiliser de vrais périphériques à condition qu'elle puisse être contrôlée à distance par un ordinateur.