

Audio/Video Fusion for Objects recognition and Robot navigation

Ryad Benosman

Institut des Systèmes Intelligents et de Robotique (UMR 7222)
Université Pierre et Marie Curie
4 Place Jussieu
75252 Paris cedex 05
France

Fusion vidéo-audio : enjeux et blocages

Ryad Benosman

Institut des Systèmes Intelligents et de Robotique (UMR 7222)
Université Pierre et Marie Curie
4 Place Jussieu
75252 Paris cedex 05
France

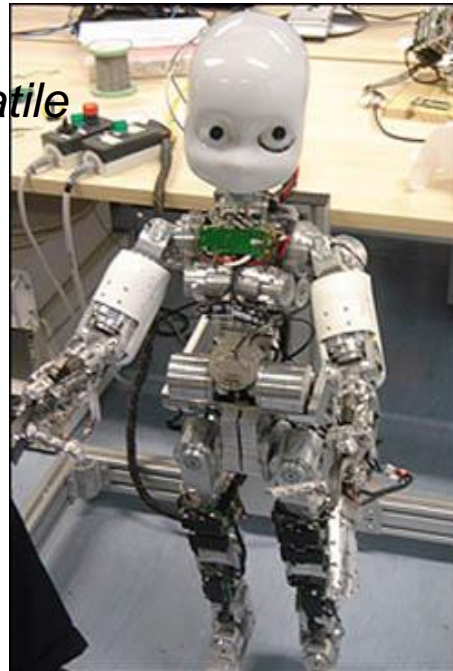
Perception Multimodale

Objectifs :

*Développer de la perception versatile
Rapide,
Economique énergétiquement*

Contraintes :

- *Temps réel*
- *Robuste*
- *Synchronisation*



ICUB

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

FASTNAV- FAST

ASAROME

QuickTime™ et un
décompresseur TIFF (non compressé)
sont requis pour visionner cette image.

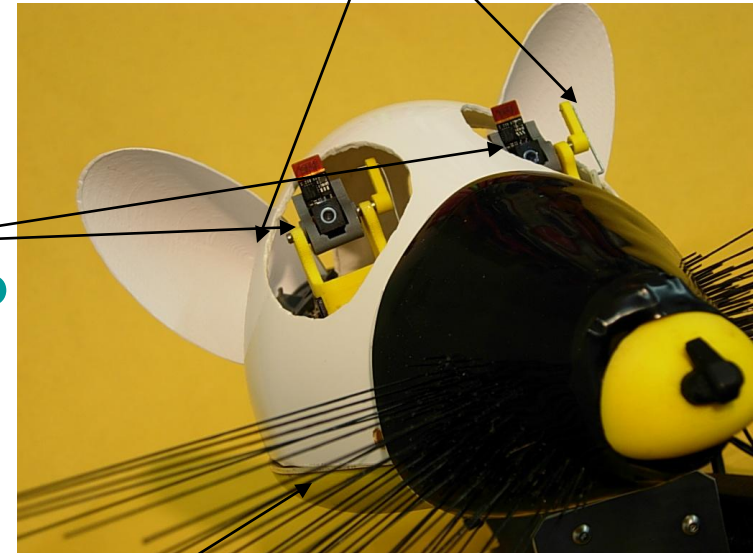
Perception Multimodale

Objectifs :

- *Développer de l'autonomie*
- *Fusion de données pour localisation, mapping,*
- *Système biomimétiques, méthodes d'ingénierie*

Vision Stereo

Audition
Binaurale



Moustaches

ICEA (FP6-IP)

Fusion de données

Schéma classique (Adelson et Bergen)

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

Primitives images : SIFT

QuickTime™ et un
d compresseur
sont requis pour visionner cette image.

QuickTime™ et un
d compresseur
sont requis pour visionner cette image.

QuickTime™ et un
d compresseur
sont requis pour visionner cette image.

Primitives images : SIFT

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

Reconnaissance d'objets :

- Appariement basé sur des **occurrences de points**
- Difficulté **d'intégrer** des **contraintes spatiales**
- Extrêmement **gourmand en mémoire et en temps de calcul**

- Base de données de **patches élémentaires**
- Décomposer l'image en patches
- Reconnaître les objets selon les **occurrences de patches**

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

Reconnaissance d'objets mobiles sonores

QuickTime™ et un
d codeur
sont requis pour visionner cette image.

- *Base de donn es d'objets sonores*
- *S quences d'objets en mouvement avec occlusions*
- *Signature sonores proches*

Primitives image

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

- **Décomposition** statique d'une image en patches.
- Classement des patches selon des critères à définir

Extraction de patches

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

QuadTree Entropique

- Découpage de l'image en sous images **équivalentes d'un point de vue entropique**
- Patches de tailles différentes
- Découpage selon la **disposition spatiale** de l'information

QuickTime™ et un
d ecompresseur
sont requis pour visionner cette image.

QuickTime™ et un
d ecompresseur
sont requis pour visionner cette image.

- D ecoupage **multi- chelle** de l'image
- Meilleure couverture de l'image
- **Stabilit  de la d ecomposition** si l'information est stable

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

- Modèle **cochléaire de Patterson** : filtres selon la bande passante
- **Modélisation prédictive des sources** sonores par des réseaux de neurones
- **Prédiction du prochain vecteur** à partir des données antérieures

Reconnaissance sans fusion

QuickTime™ et un décompresseur sont requis pour visionner cette image.

Reconnaissance avec fusion

QuickTime™ et un décompresseur sont requis pour visionner cette image.

GD1 : Erreur globale
GD2 : Majorité simple

Deux types de décisions locales sont effectuées.

- Cumuler les erreurs audio et vidéo locales : **erreur locale commune**
- Calculer **deux erreurs indépendantes**, effectuer plus tard une décision globale.

La décision globale étant bien sur d'affecter une classe d'objets à l'ensemble d'une séquence à partir des décisions locales.

Fusion : limitations

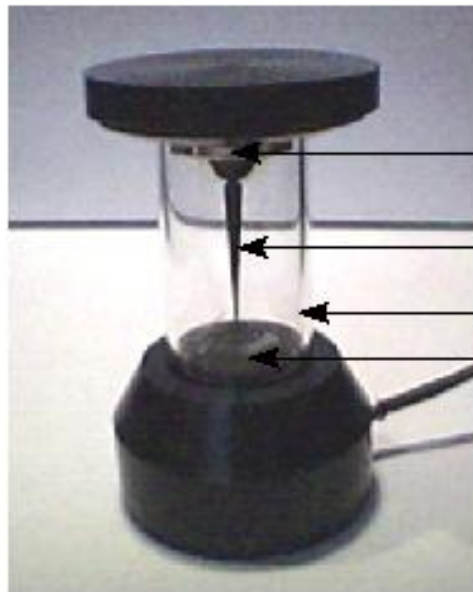
QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

2D

1D

- Incompatibilité de **format**
- Incompatibilité de **fréquences**
- Incompatibilité de **périodes d'observations**

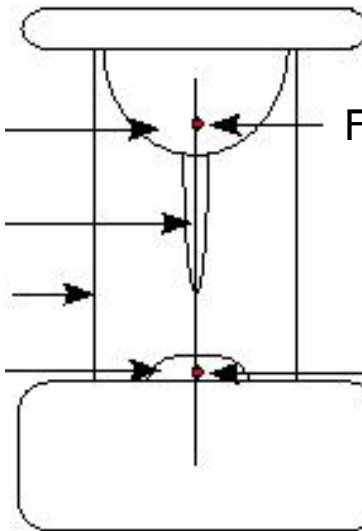


miroir

aiguille

Tube

caméra

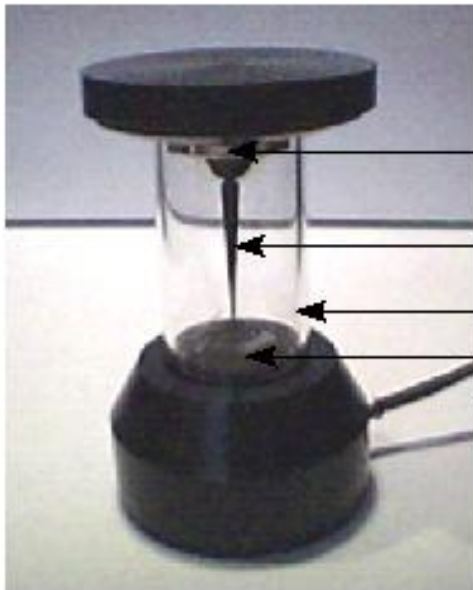


Foyer miroir

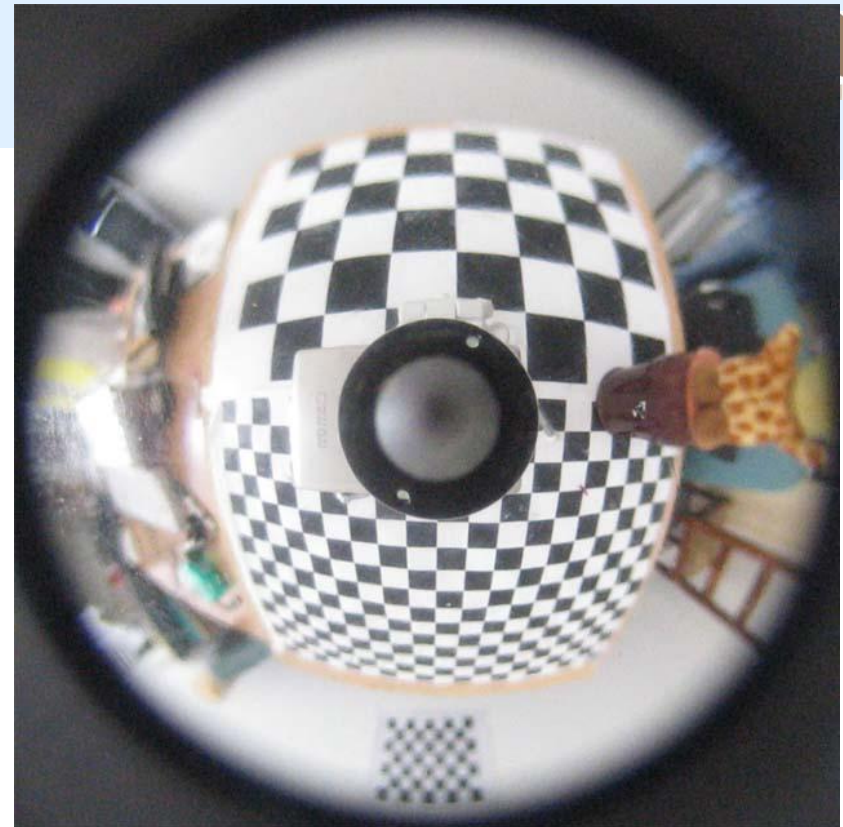
Foyer caméra

Capteurs Catadioptriques

- Association miroir - caméra
- Surface réfléchissante non linéaire
- Acquisition rapide : temps réel



miroir
aiguille
Tube
caméra

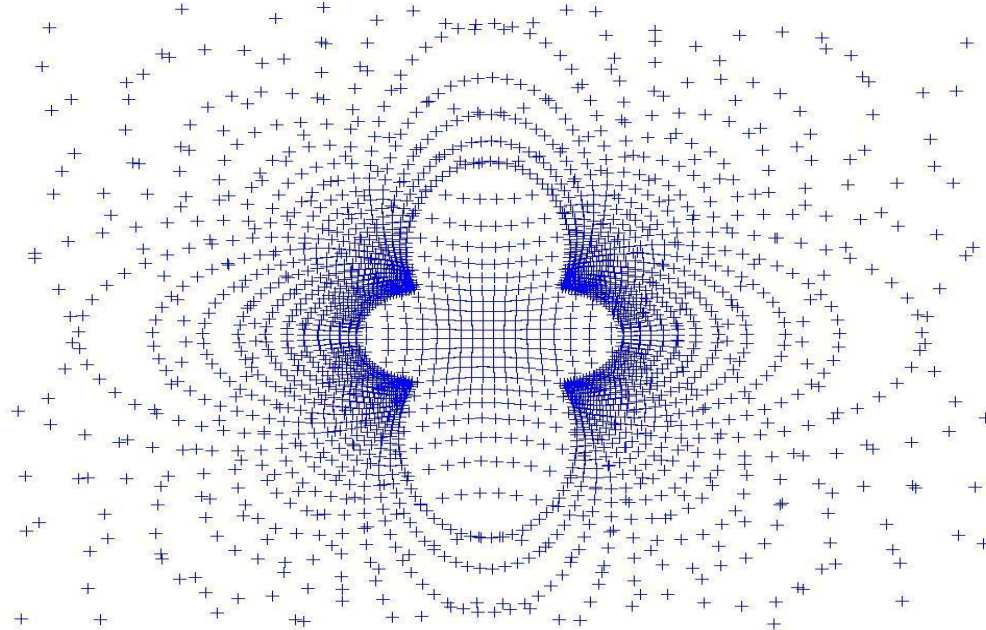


Capteurs Catadioptriques

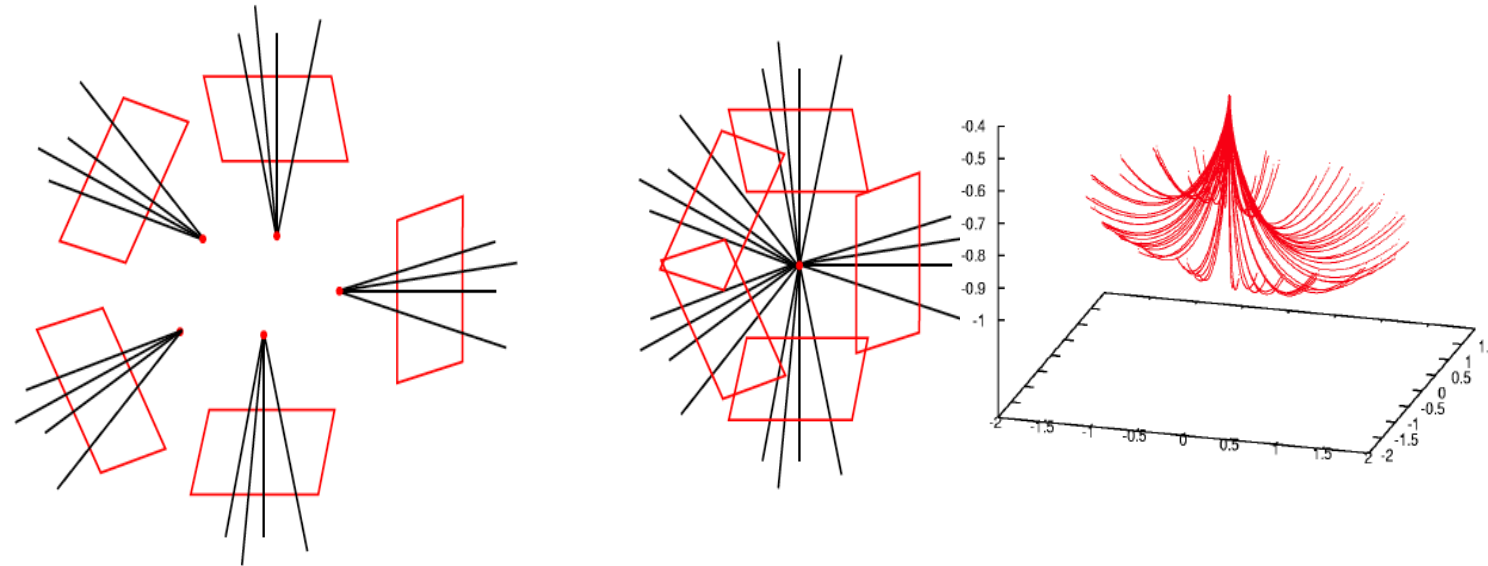
- Association miroir - caméra
- Surface réfléchissante non linéaire
- Acquisition rapide : temps réel

ABSENCE D'UN MODÈLE DE TRAITEMENT DES IMAGES

Limitations du modèle central



- Condition **hypothétique**
- Fortes contraintes sur la géométrie des capteurs
- Couverture de la scène non adéquate avec les applications



QuickTime™ et un
d compresseur TIFF (non compress )
sont requis pour visionner cette image.

- **Capteurs catadioptriques : G n ralisation des capteurs de vision**
- Construction de n'importe quel capteur (champ de vue et r solution)
- **Absence de primitives visuelles.**

QuickTime™ et un
d compresseur
sont requis pour visionner cette image.

QuickTime™ et un
d compresseur
sont requis pour visionner cette image.

QuickTime™ et un
d compresseur
sont requis pour visionner cette image.

QuickTime™ et un
d compresseur
sont requis pour visionner cette image.

- **Systemes biologiques** plus **performants** que n'importe quel systeme artificiel
- Unicit  du langage : **spikes**
- Importance du **temps et de la dynamique de l'acquisition**

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

QuickTime™ et un
décompresseur
sont requis pour visionner cette image.

- **Systemes biologiques** plus **performants** que n'importe quel système artificiel
- Unicité du langage : **spikes**
- Importance du **temps et de la dynamique de l'acquisition**

Repenser les capteurs