LABORATOIRE DES SCIENCES DU NUMÉRIQUE DE NANTES

PROJET FINANCÉ PAR L'ANR - PROJECT FUNDED BY THE ANR

Influence of viewpoint on visual saliency models for volumetric content

Mona Abid, Matthieu Perreira Da Silva and Patrick Le Callet

IPI - Image Perception Interaction 25.09.2019

ICIP 2019: International Conference on Image Processing

ICIP 2019

IEEE International Conference on Image Processing

UNIVERSITÉ DE NANTES

Nantes Métropole

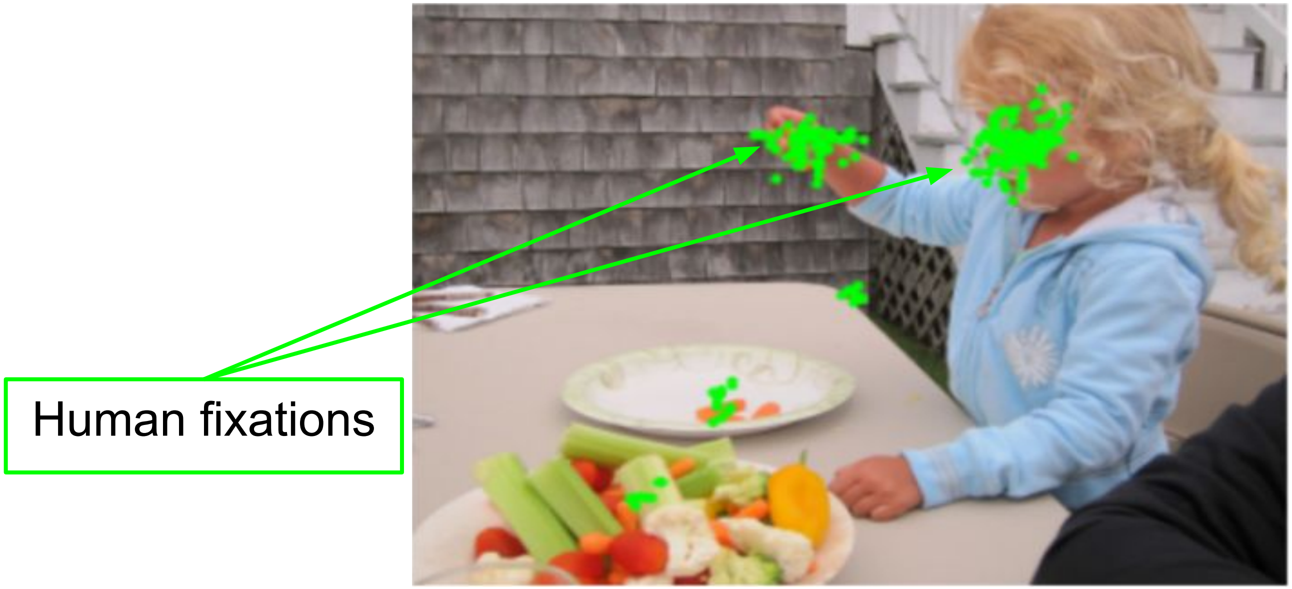
Context

Visual attention and saliency usage

Look at the image

Context

Visual attention and saliency usage



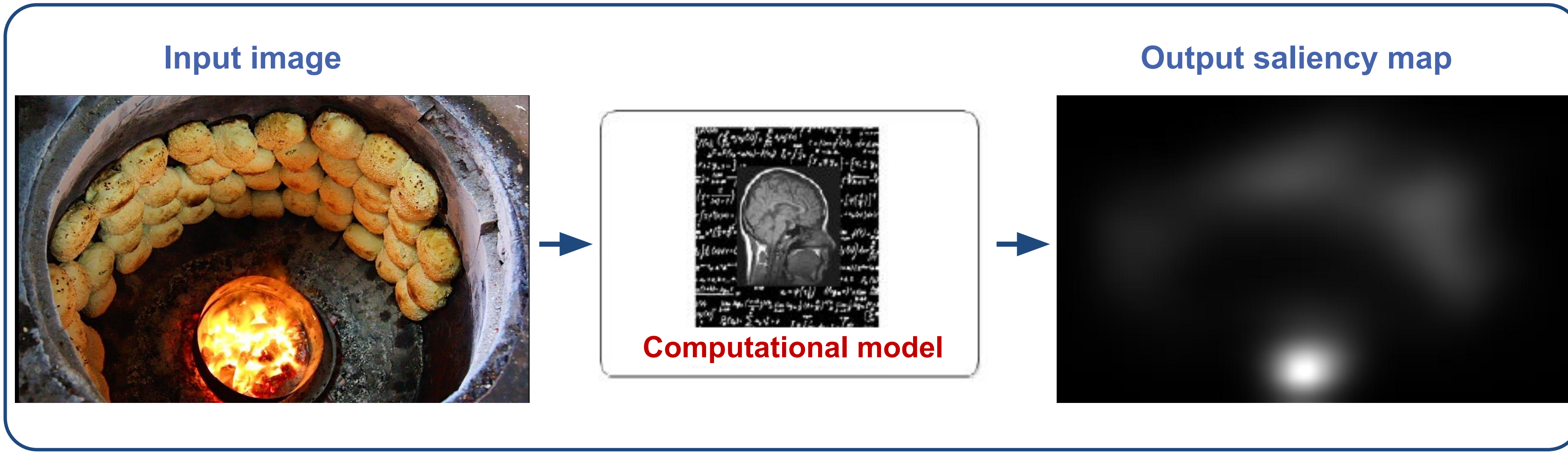
Utility

Improve the user's quality of experience

Context

<u>How to get saliency information?</u>

Computational models aim to predict where we look within a scene.



Challenge

Visualisation conditions change according to the context (3D context, immersive context, etc).

Problem statement

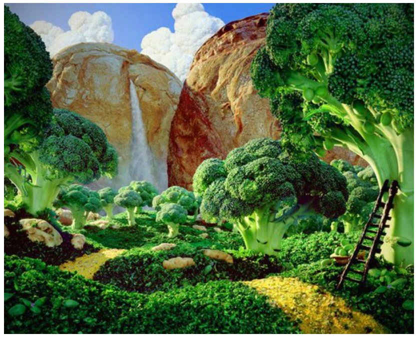
Visual saliency in 3D immersive environment

- Immersive context is different from 2D natural scenes

1. Content characteristics (material, texture, etc)

From: **Natural scenes**

To: **Computer Generated (CG) scenes**



Problem statement

Visual saliency in 3D immersive environment

- Immersive context is different from 2D natural scenes

2. Interaction is possible (forward, backward, etc)

➡ No existing recommendations

- Dataset availability
- Experimental protocol to conduct
- Number of observers to consider

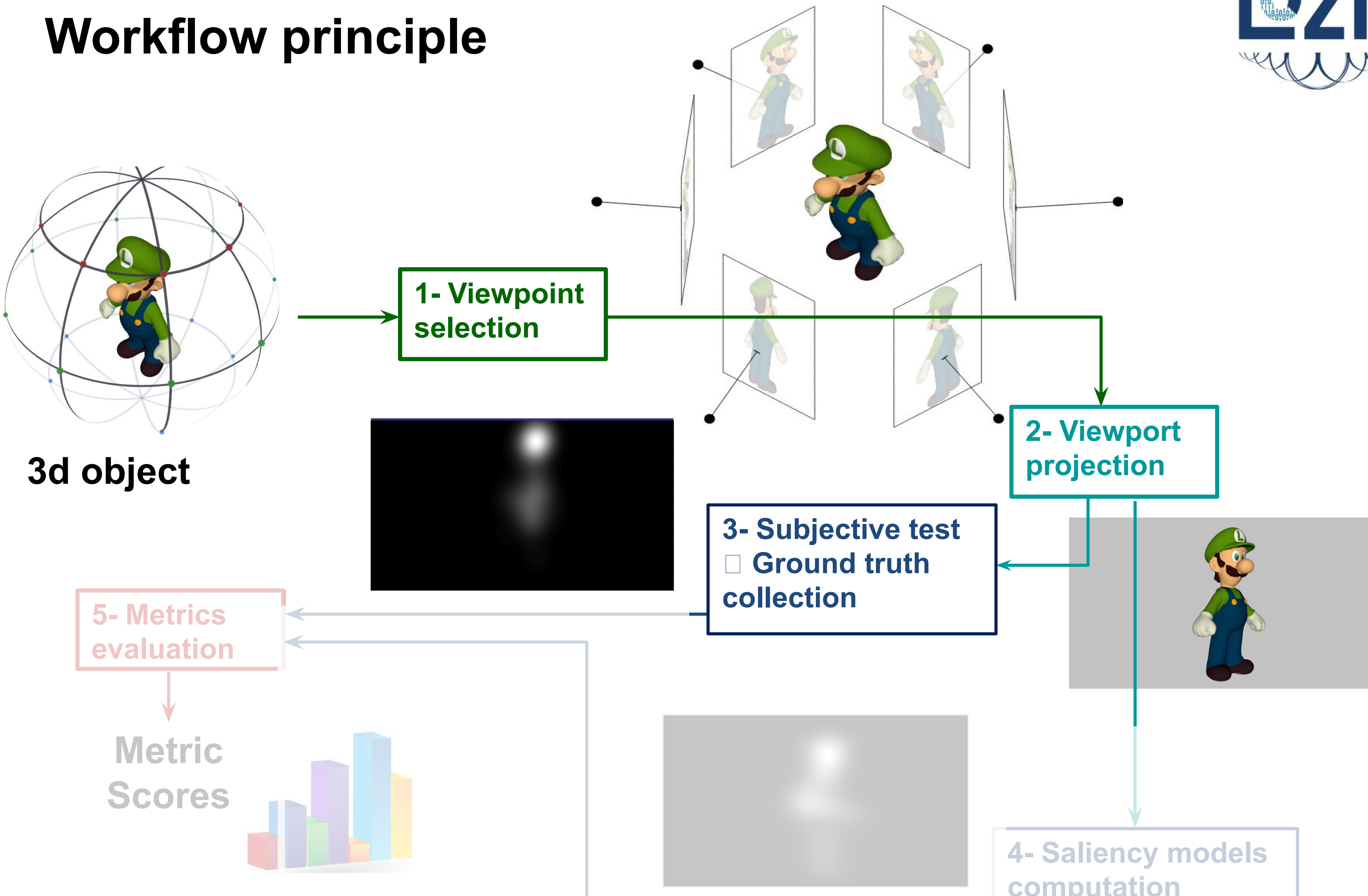
Open questions

- Effect of usage of different content type (eg. 3D contents) on gaze
- Usage of non-conventional displays to explore contents

Outline

- How well current saliency models perform on computer generated contents?
- Impact of viewing distance on visual saliency models in immersive context
- Conclusion and ongoing work

Workflow principle



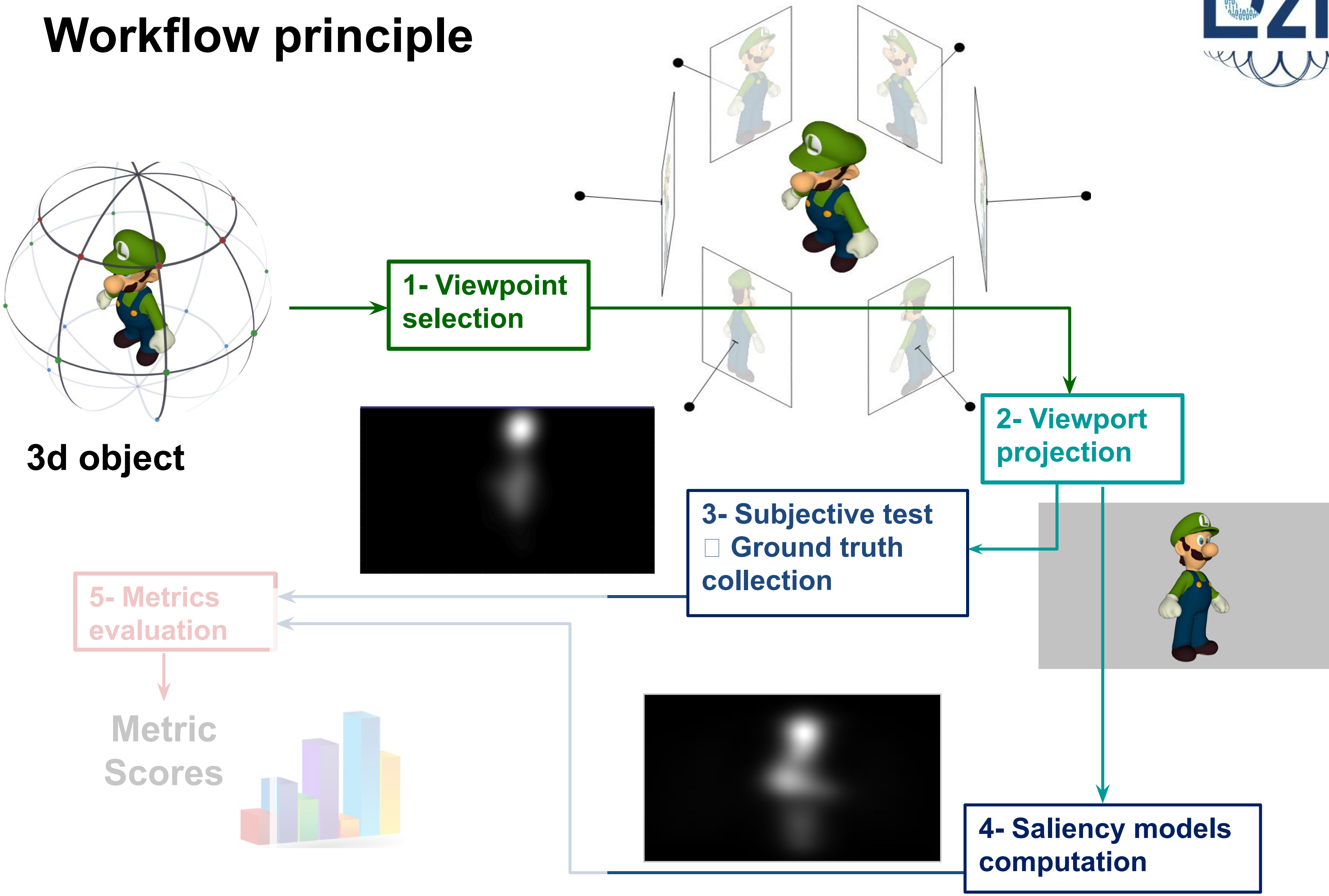
Psycho-visual subjective experiment

- Rendering:
  - Fixed viewpoint
  - 2D screen
  - 30 observers with normal/corrected-to-normal vision
  - 3 seconds per stimulus

Example of used CG contents

1. How well current saliency models perform on computer generated contents?

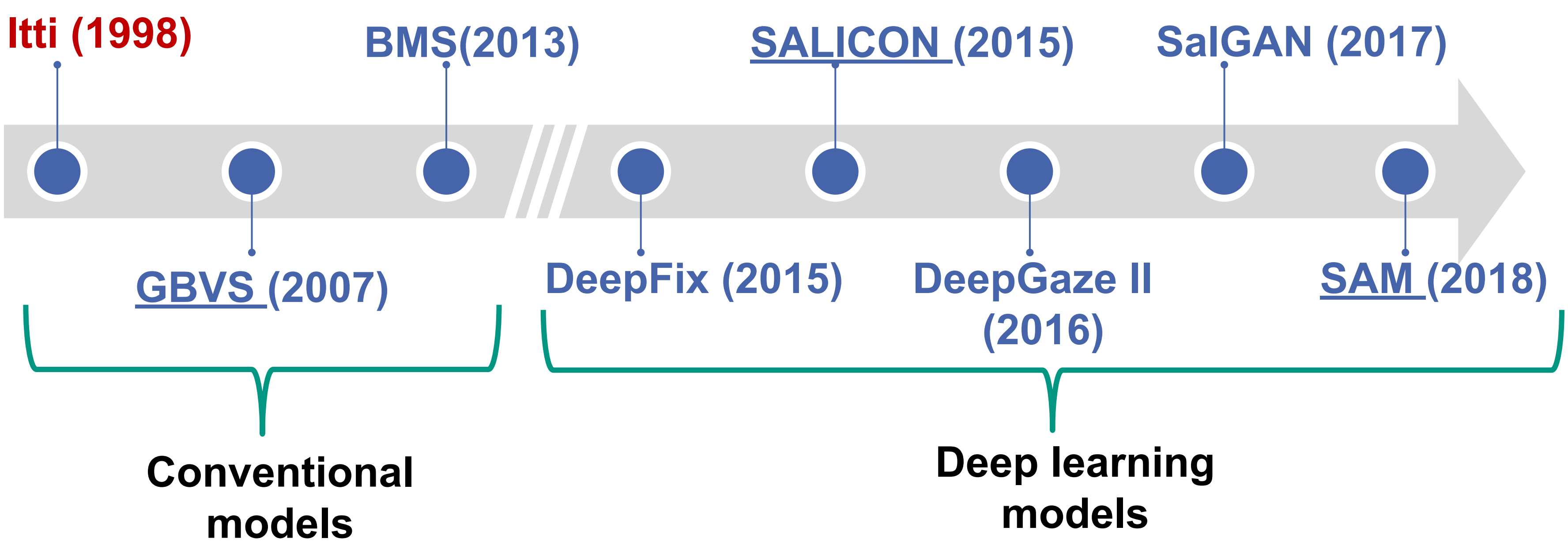
Workflow principle



Overview of visual attention models

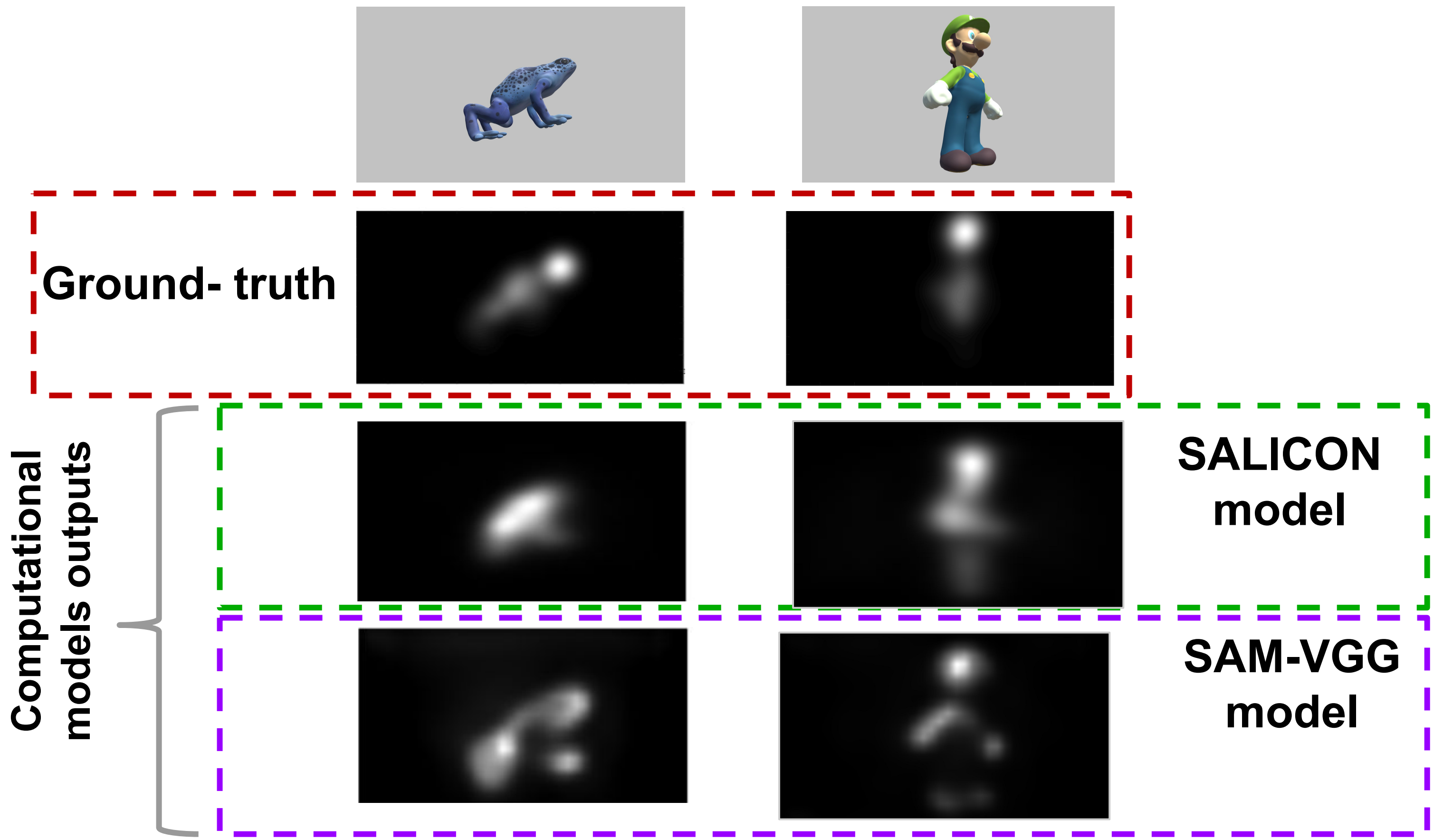
2D visual attention prediction based on:

- **Conventional approaches:** low features extraction: color, intensity, orientation, etc.
- **Deep learning techniques:** using neural networks.

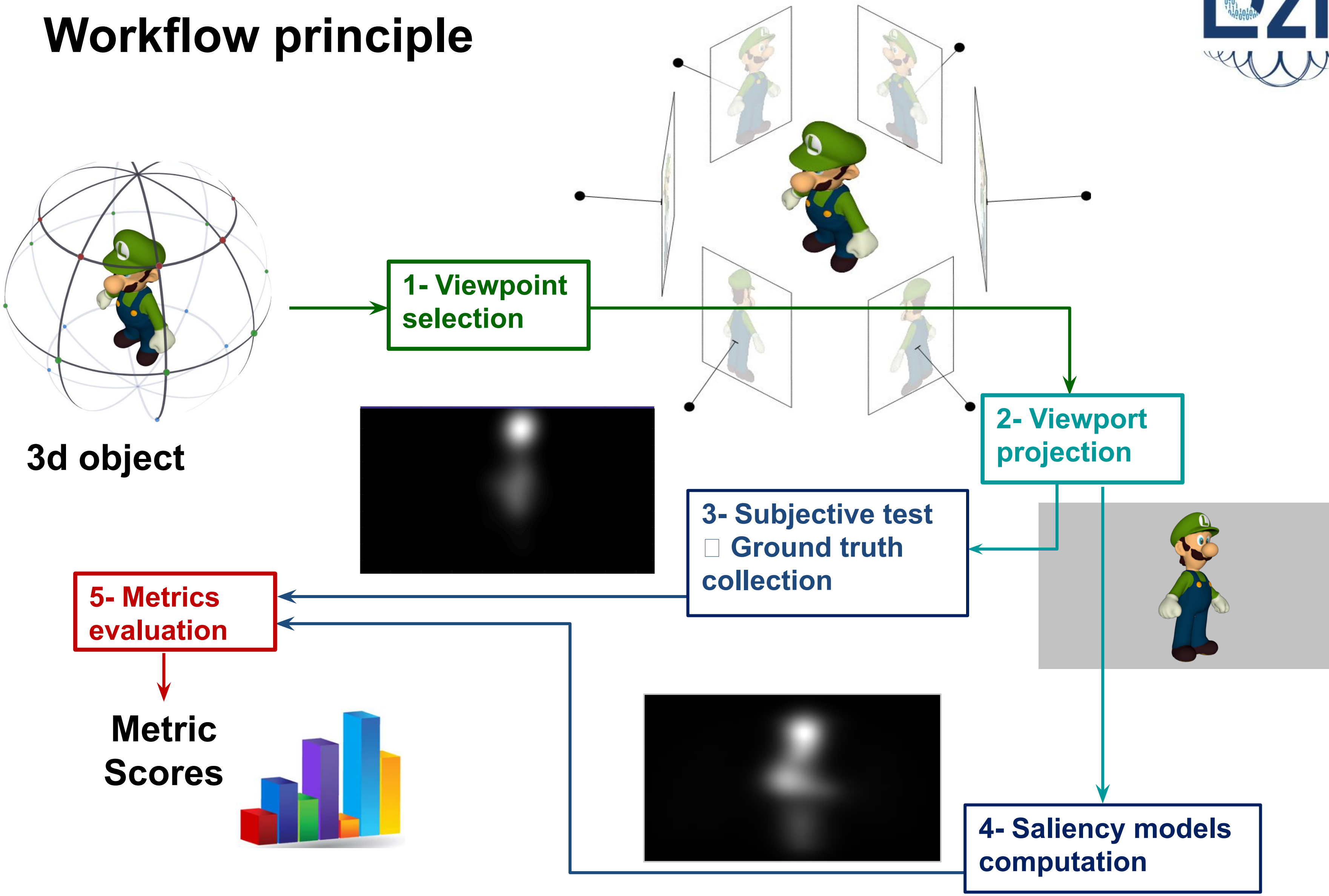


Application on computer generated contents

- Ground-truth dataset collection
- Saliency models computational



Workflow principle



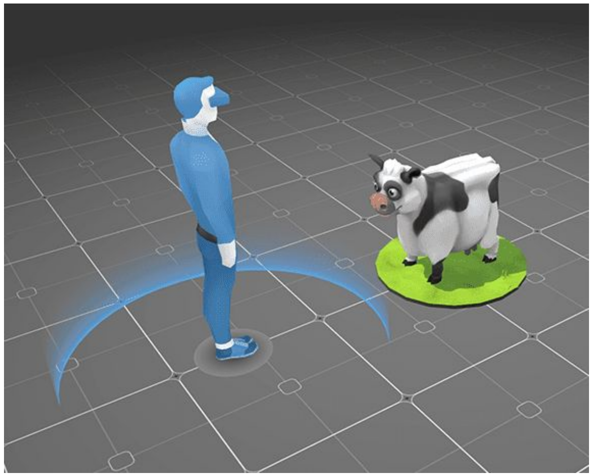
Results: Metric measurements

Saliency model	Metrics	Mean
SALICON	NSS ↑	**1.06**
	KLD ↓	**0.54**
SAM-Vgg	NSS ↑	0.59
	KLD ↓	0.94
SAM-Resnet	NSS ↑	0.70
	KLD ↓	0.87
GBVS	NSS ↑	0.61
	KLD ↓	1.05

Conclusion

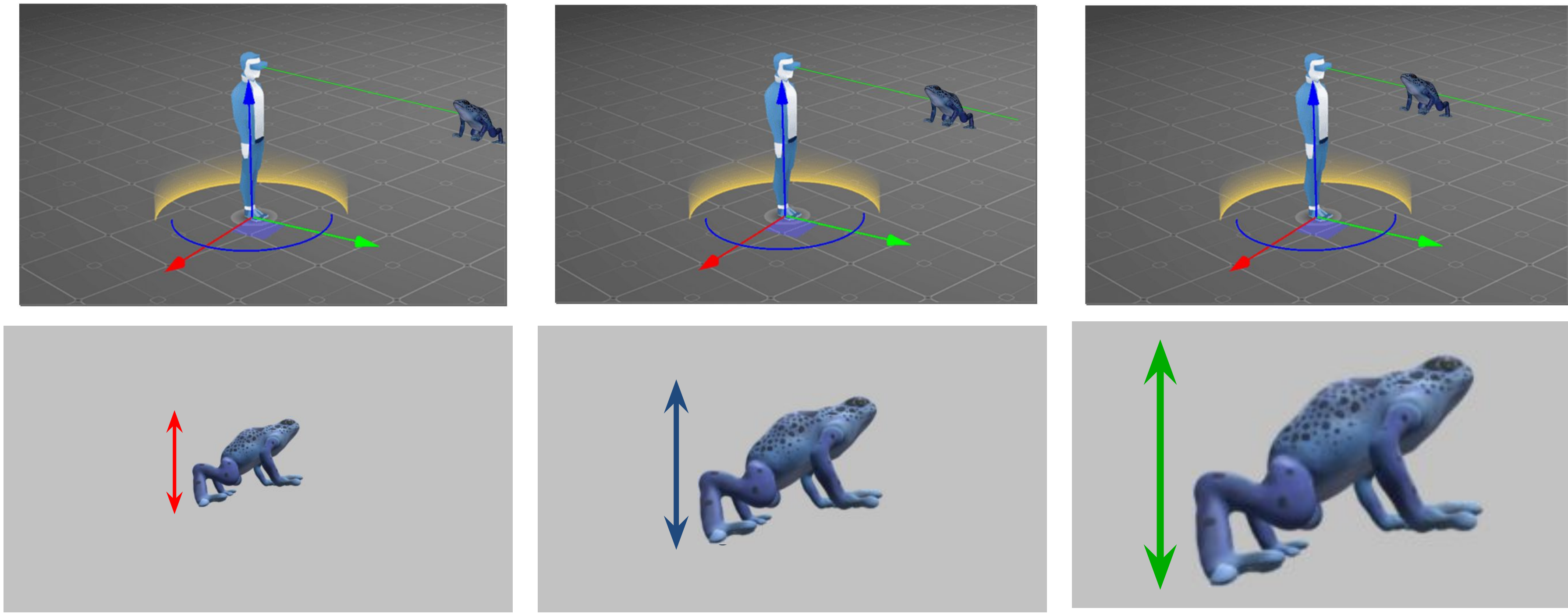
Among evaluated saliency models, **Salicon** seems to be the **most suitable model** when applied on **computer generated contents.**

Influence of the viewing distance on visual saliency models in immersive context



Quantified viewing conditions: 3 viewing distances for each CG object

Stimuli generation - protocol design



Small scale | Medium scale | Big scale

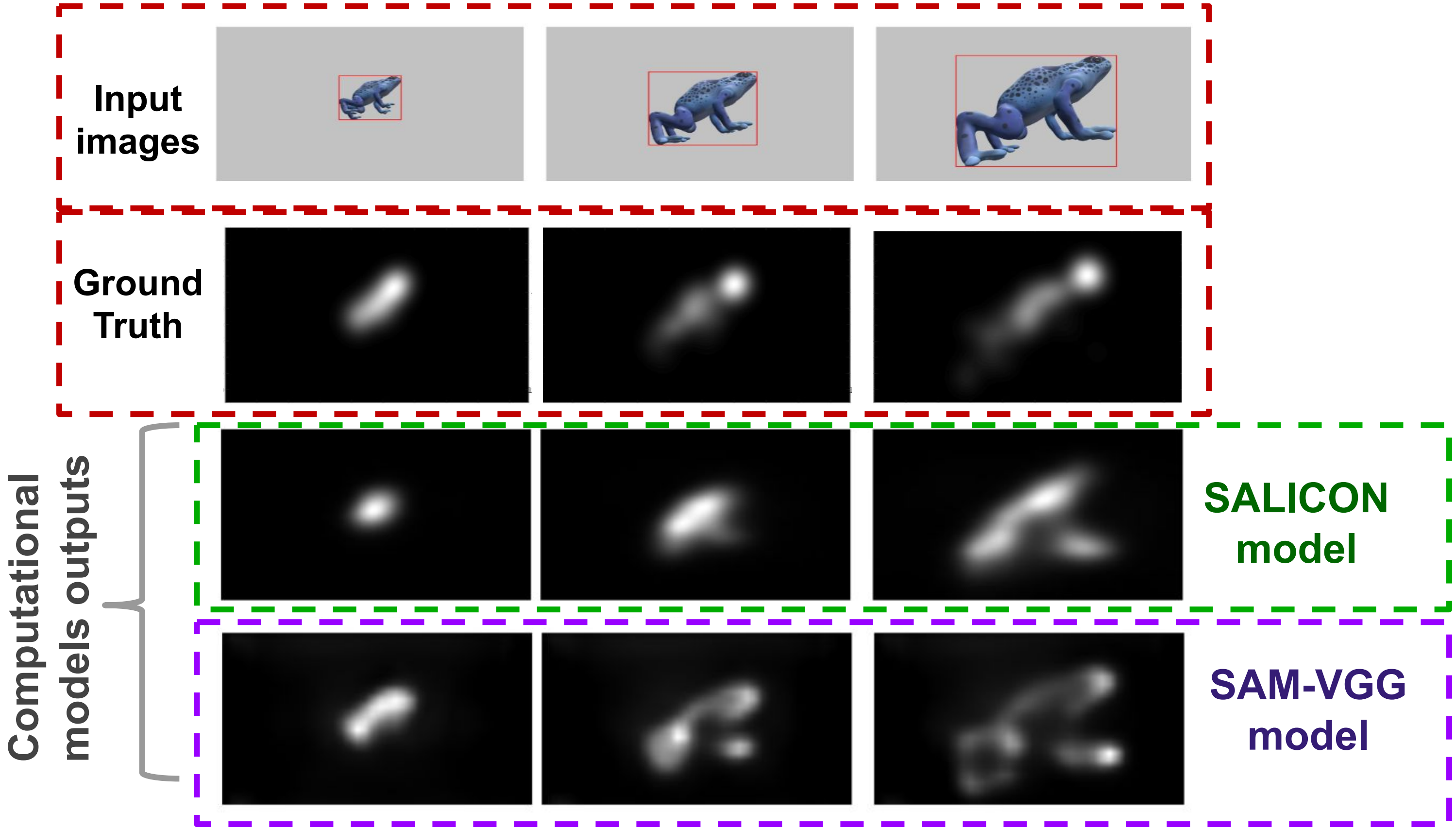
3 variantes for each CG object :

- **Size** variation
- **Level of Detail (LoD)** variation

24 objects x 3 sizes
72 stimuli used for the experiment

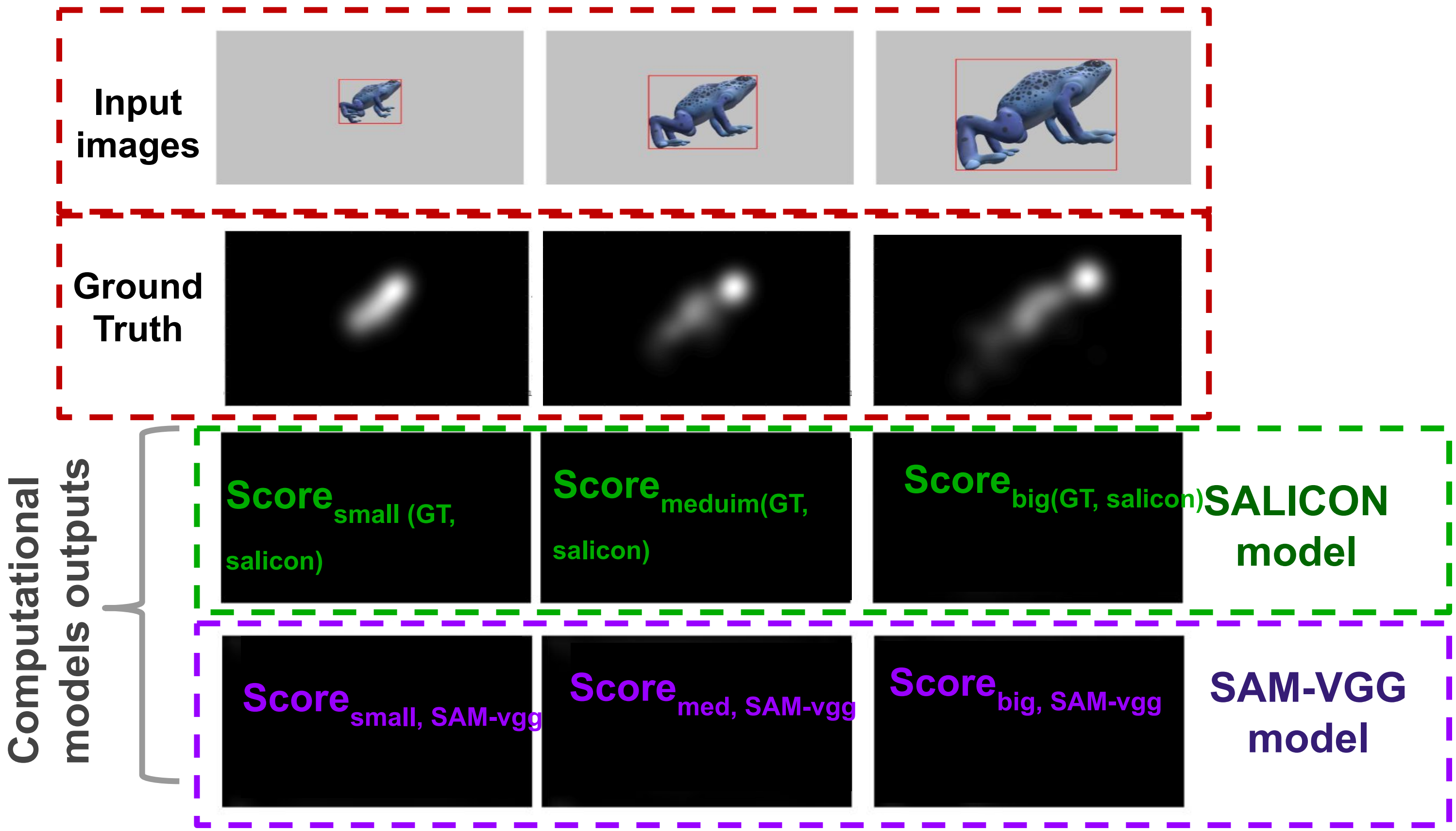
Data processing

- ❑ Ground-truth gaze data collection
- ❑ Saliency models computation



Data processing

Metrics evaluation on the bounding box (useful information)



?

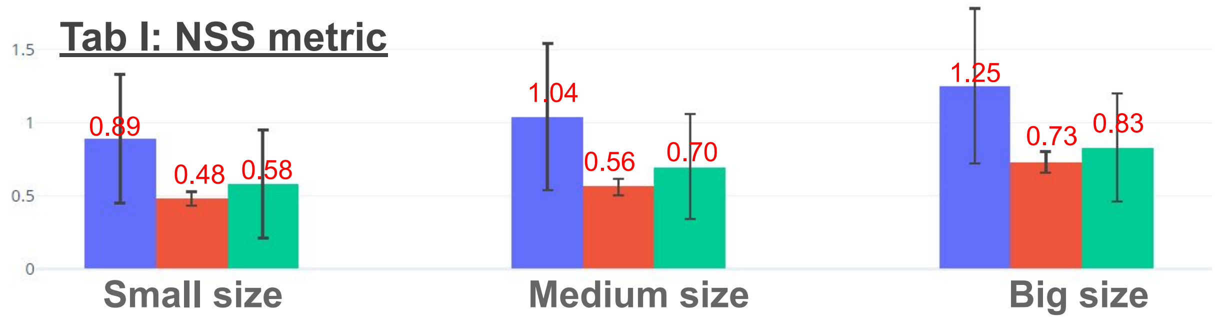
For a given model, how significant metric scores vary when varying the viewing distance ?

Saliency models

- Salicon
- SAM-Vgg
- SAM-Resnet

Mean and standard deviation values on all objects

Tab I: NSS metric



Tab II: KLD metric

Yes, but ...

For a given model, metrics scores vary from viewing condition to another.

For a given model, how significant metric scores vary when varying the viewing distance ? → **ANOVA**

Saliency model	Metrics	p-value	Statistical significance
Salicon	NSS	0.0461	✓
	KLD	0.0134	✓
SAM-Vgg	NSS	0.0062	✓
	KLD	0.0011	✓
SAM-Resnet	NSS	0.0670	✗
	KLD	0.0008	✓

If p-value < 0.05, metric scores of the 3 variants (3 content sizes) have statistically significant differences for a given model.

Conclusion

The content size has significant impact on models performances.

Conclusion

- **Saliency models** have relatively low performances when applied on **CG contents** compared to **natural 2D images**.

- Among these models, **Salicon** seems to be the **most suitable model** when applied on **computer generated contents**.

- Overall, computational models are **not robust** to **viewing distance change**.

On going work

- **Integrate** the **viewing distance** information as a **parameter** in computational models to **improve their performances.**

- Consider **fine-tuning** deep saliency models (training on CG contents).

Thank you for your attention!

Perceptual Levels of Detail for Interactive and Immersive Remote Visualization of Complex 3D Scenes

PISCo

UNIVERSITÉ DE NANTES

Nantes Métropole

Contact: mona.abid@univ-nantes.fr

AGENCE NATIONALE DE LA RECHERCHE

This work was funded by the French National Research Agency as part of **ANR-PISCo** project