MAR Visualization Requirements for AR based Training

Gerard J. Kim, Korea University

2019 SC 24 WG 9 Presentation (Jan. 23, 2019)

Information displayed through MAR?

- Content itself
  - Associate target object (in real or virtual environment) with
  - Augmentation object (which is virtual or captured/reconstructed from the real)
- Expectation
  - Augmented and composited display will be “visible”, “natural”, “seamless” and “consistent”
  - Is it?
- MAR is used in a dynamic condition
  - User/target is moving (mobile)
  - User/target is in different environments / timings
    - Under different lighting condition
  - This not really part of the “content”

Augmented training

- For whom?
  - Trainee – to apply and learn first hand, …
  - Trainer – to teach, instruct, and communicate better, …
  - Manager – to evaluate, …
  - Spectator – to understand, empathize and experience (learn second hand)

Augmented training

- What
  - Objects and Tools
  - Directions / Motions
  - Instructions/Strategies
- When
  - During play/training
  - Live vs. Recorded

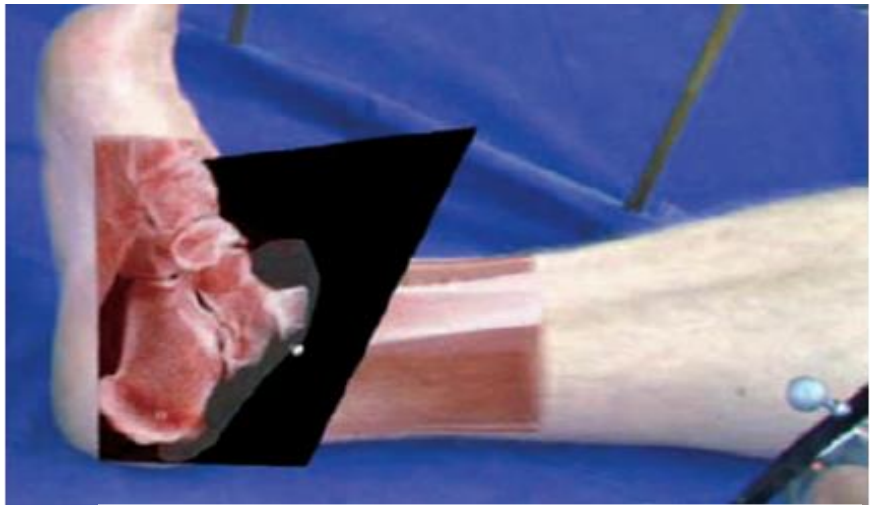
MAR Viz. Requirement 1: Naturalness

- The user is not puzzled about the imagery - it is not too much deviant from the mental model of the user's understanding of the world
  - 3D alignment: augmentation should be registered into the 3D space in a "physically plausible" way and/or provide depth cues
  - Realistic – The augmentation should look as realistic and harmonious with the real world as possible

Assumption: Spatially Registered AR

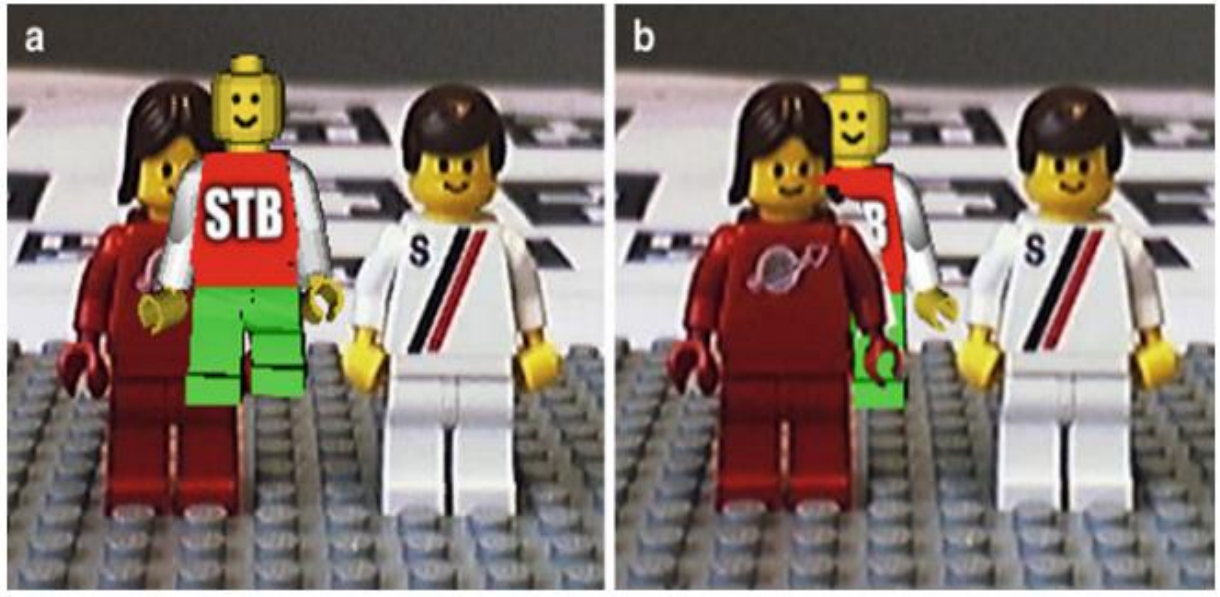
3D DEPTH CUES

Transparent overlay of 3D structures
[Azuma et al.]



3D cutting planes [Navab et al.]

(consider shadows, size, motion, ground, …)



Phantom objects for physical plausibility/occlusion
(3D model of real objects)
[Kalkofen et al.]

VISUAL REALISM / LIGHTING

Supan et al.

MR360: Mixed Reality Rendering for 360° Panoramic Videos

Victoria University of Wellington
Taehyun Rhee, Lohit Petikam, Benjamin Allen, Andrew Chalmers

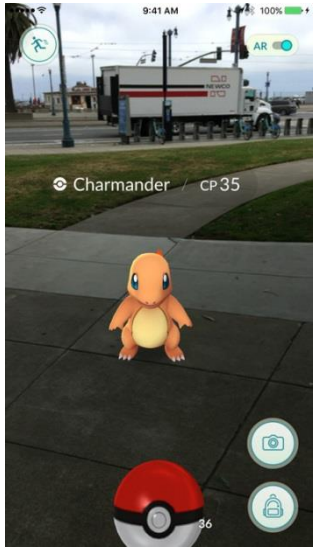
TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI
VICTORIA
UNIVERSITY OF WELLINGTON

Rhee et al.

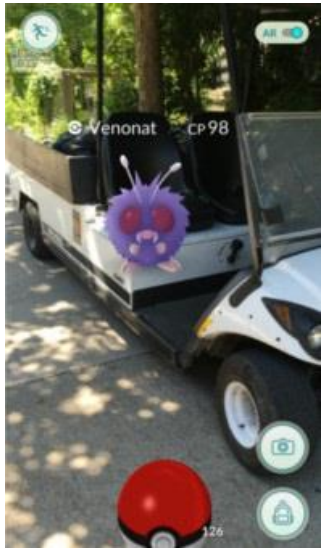
PHYSICAL PLAUSIBILITY

vs.

Avatar adaptation to remote environment [Jo and Kim et al.]



vs.



Floating Pokemons

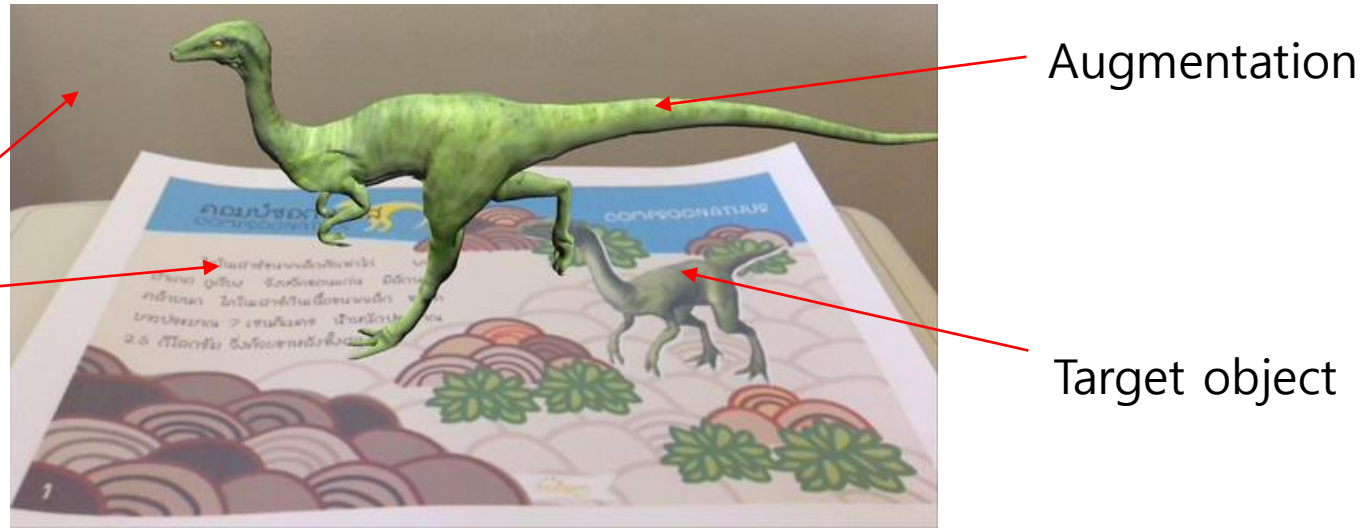
vs.

Conspicuous but unnatural?

Natural but not so outstanding

MAR Viz. Requirement 2: Visibility

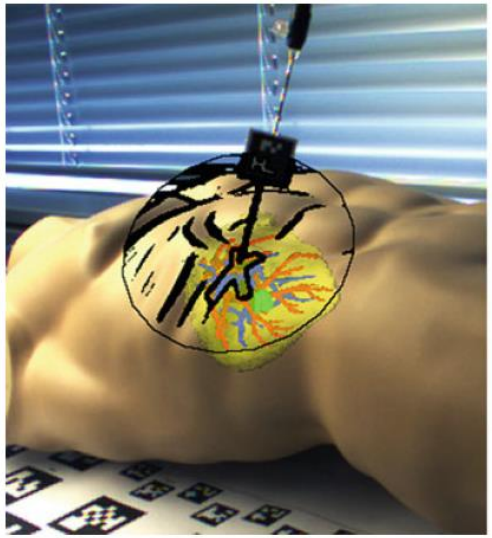
- The visibility of information should be maximized
- There are three types of information to consider



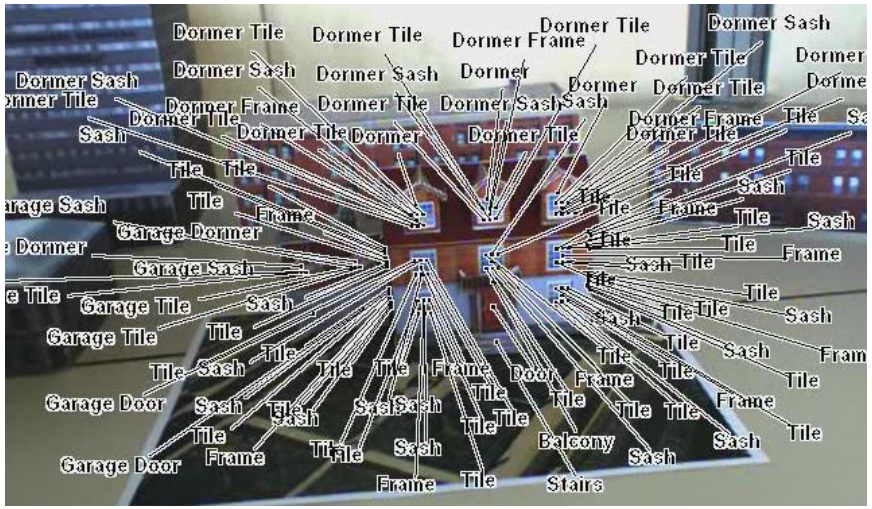
- There can be trade-offs between Maximizing Visibilities among the three

Assumption: Spatially Registered AR

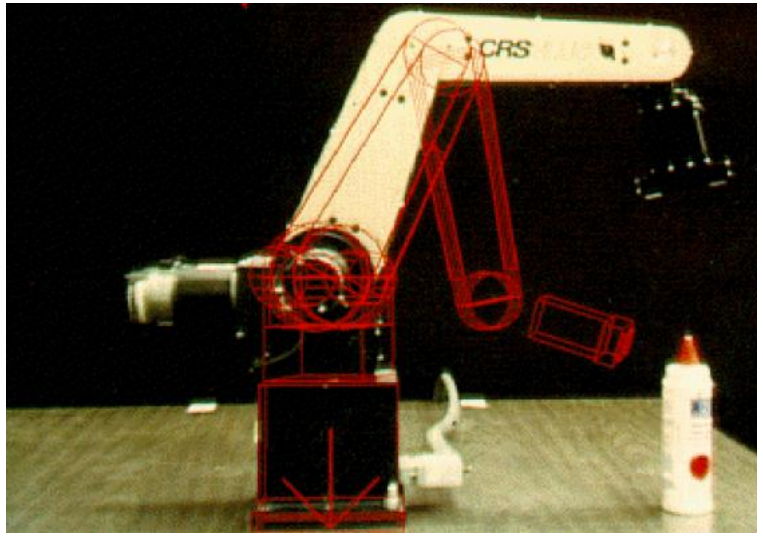
Target Object (Enhancement)



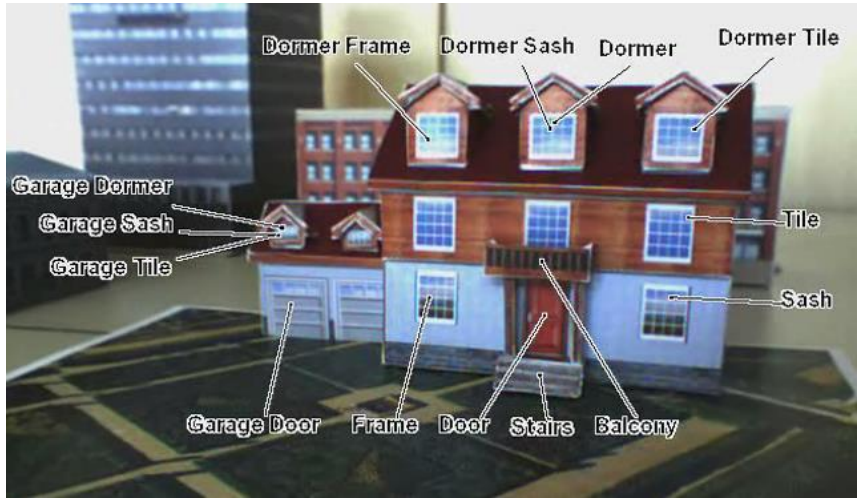
Emphasis with edges
[Avery et al.]



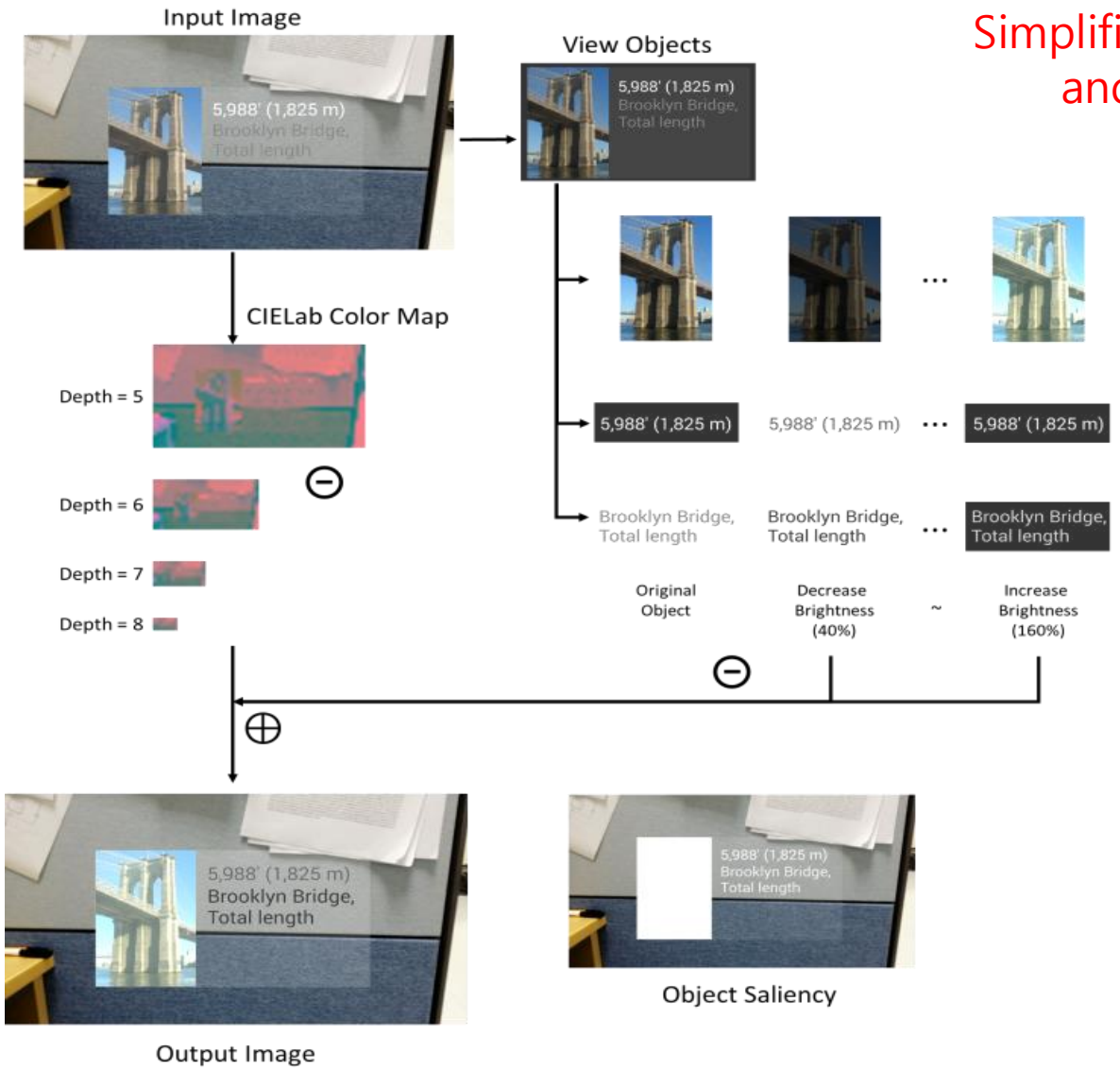
Filtering
[Tanzgem et al.]



Motion profile
[Milgram et al.]

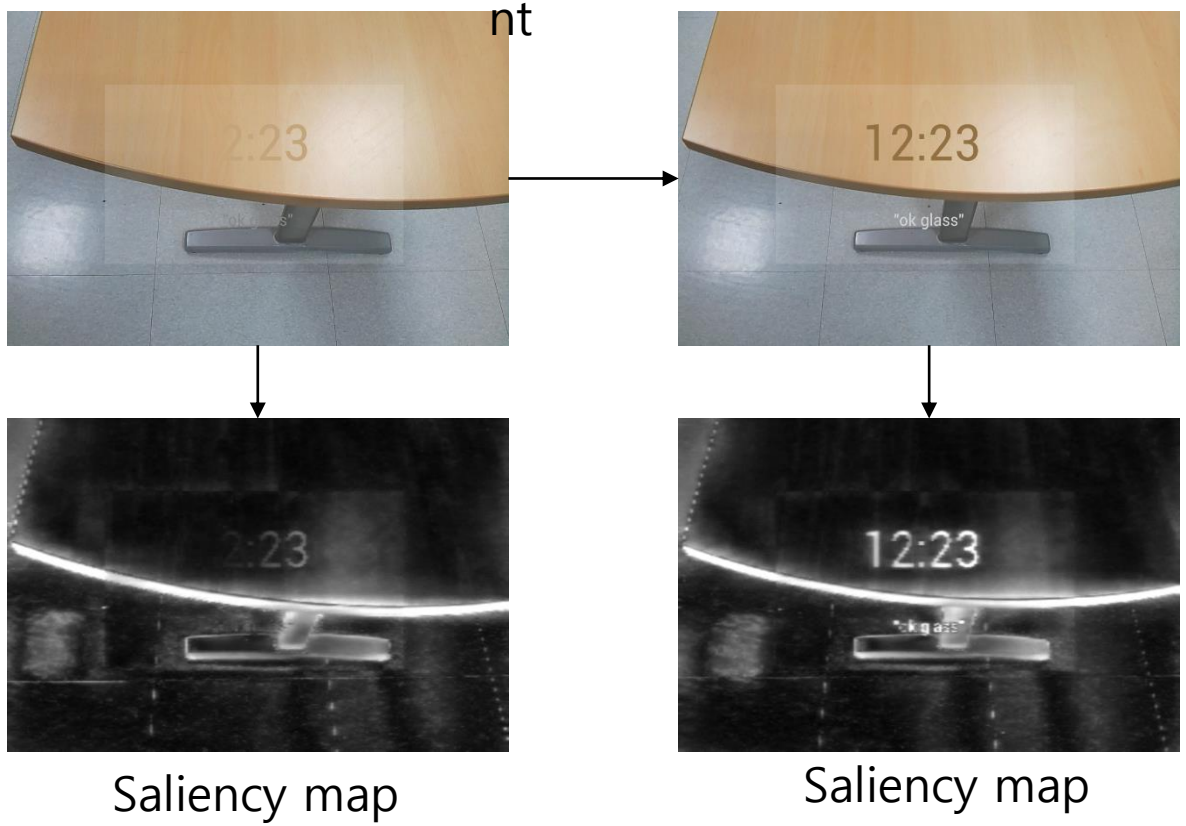


Clear Augmentation
Saliency based Method [Ahn and Kim]



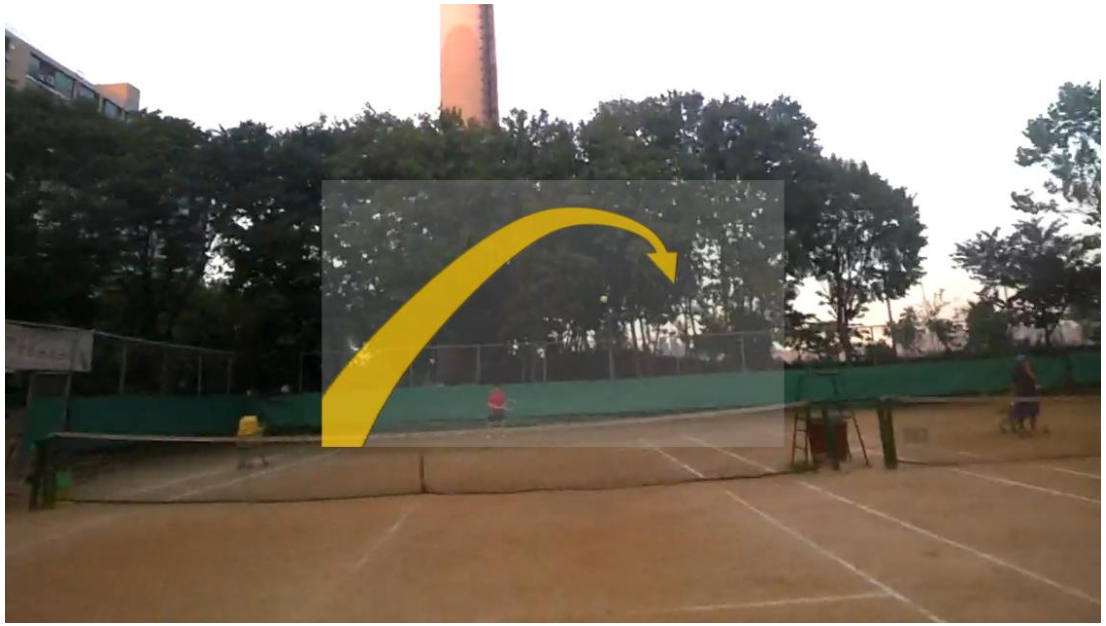
Simplified/Real time and Mobile

Adjustment



MULTIMODALITY [KIM AND KIM]

Fast paced guidance



Visual, Aural, Visual + Aural ?

movie

Augmentation and Target Object

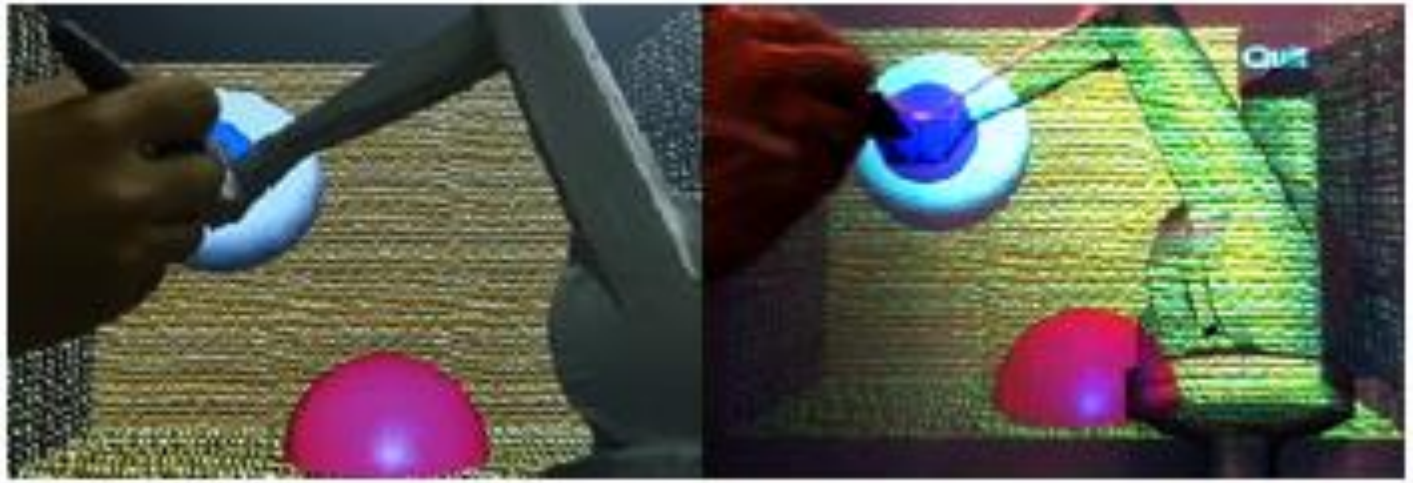
Saliency of occlude and occludee [Sandor et al.]

Transparency adjusted by saliency [Zollman et al.]

Is it possible to make both target and augmentation clearly visible at the same time?

Context / Neighborhood

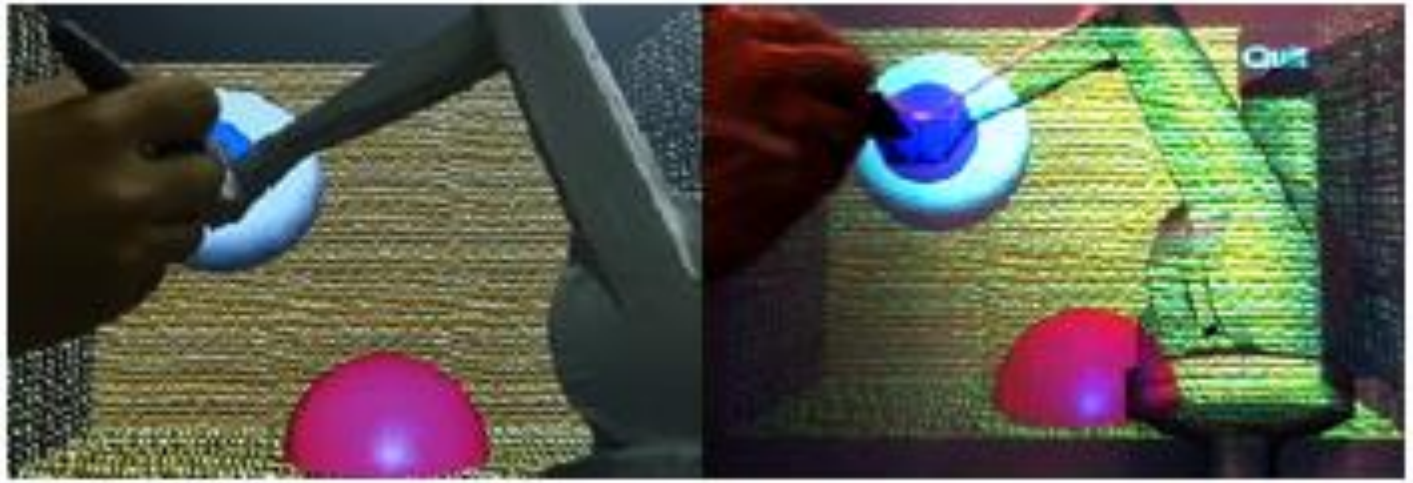
Warp to show more
Sandor et al.



Remove to show more
Inami et al.

Context / Neighborhood

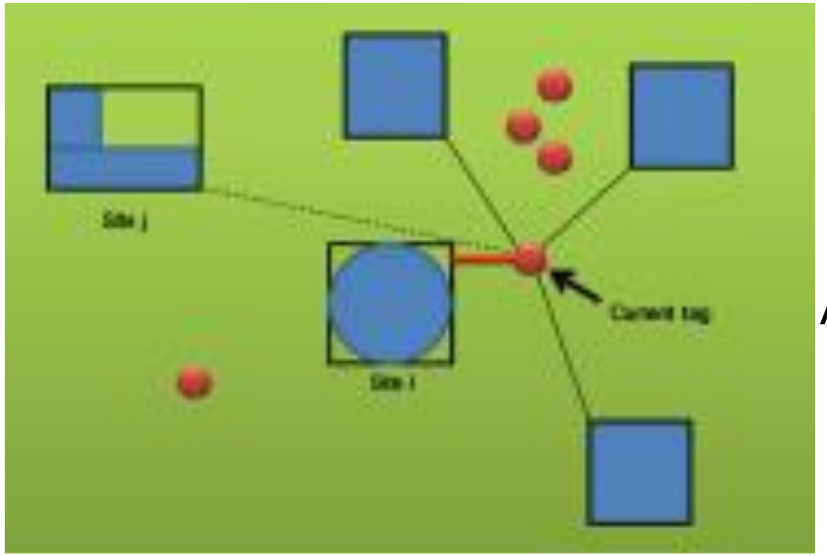
Warp to show more
Sandor et al.



Remove to show more
Inami et al.

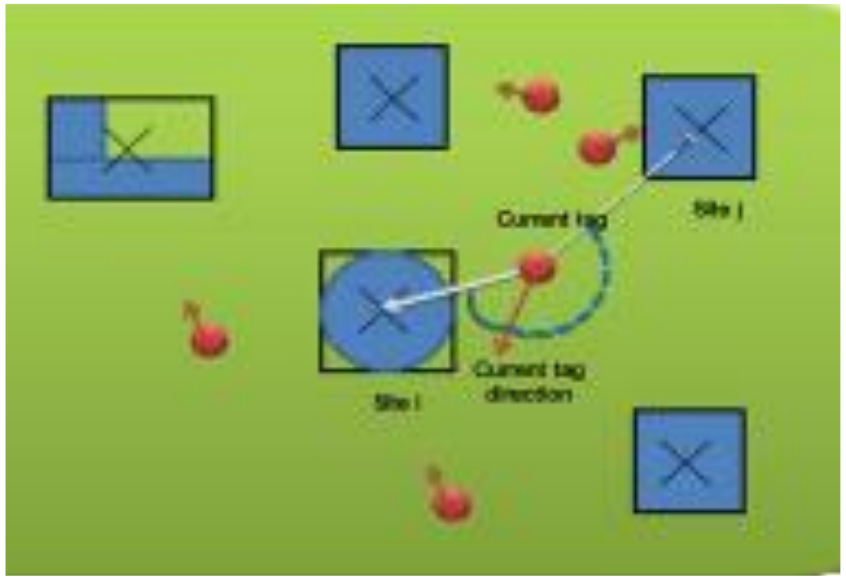
Context / Neighborhood

Crowded tag problem



Organize to show more
[Choi and Kim]

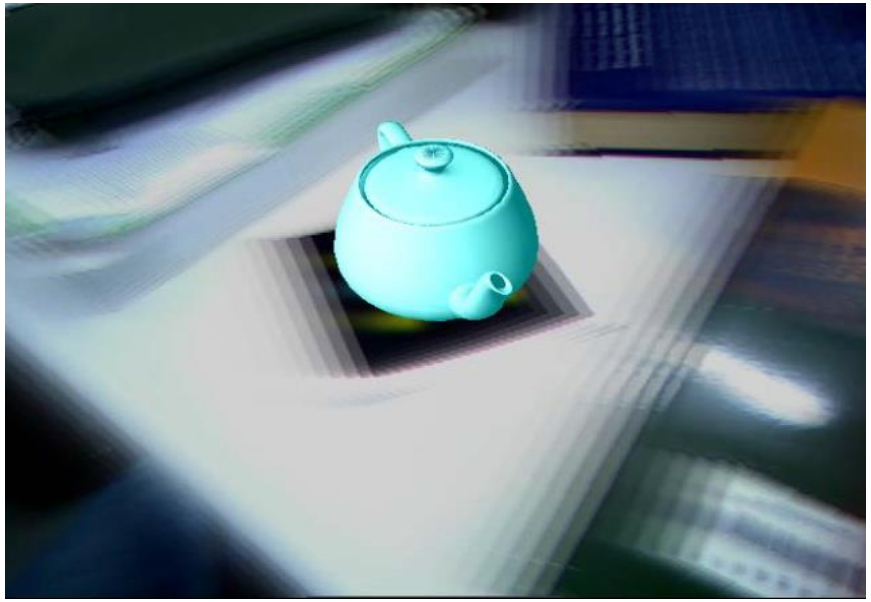
Associate Geo-tags to a group
By location/direction similarity



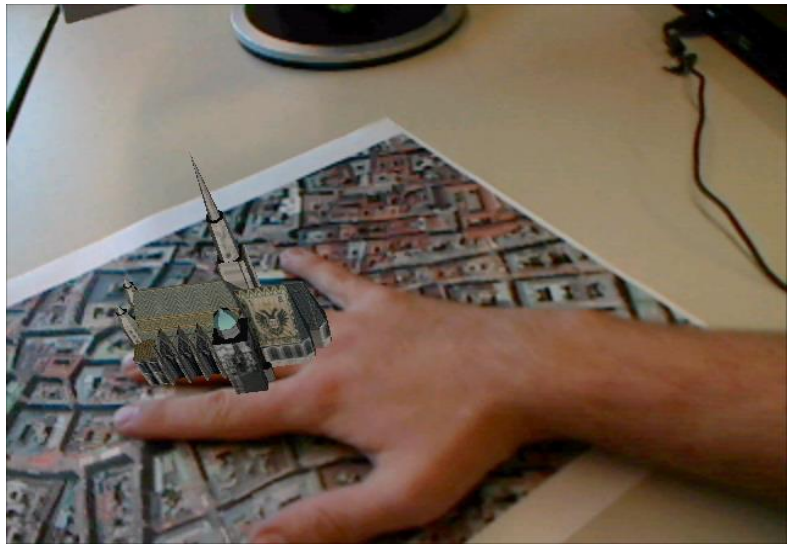
MAR Viz. Requirement 3: Stability and Persistence

- Augmentation should be stable and comfortable to View/interact
- Once augmented in certain form, it should not disappear nor change without any meaningful reason

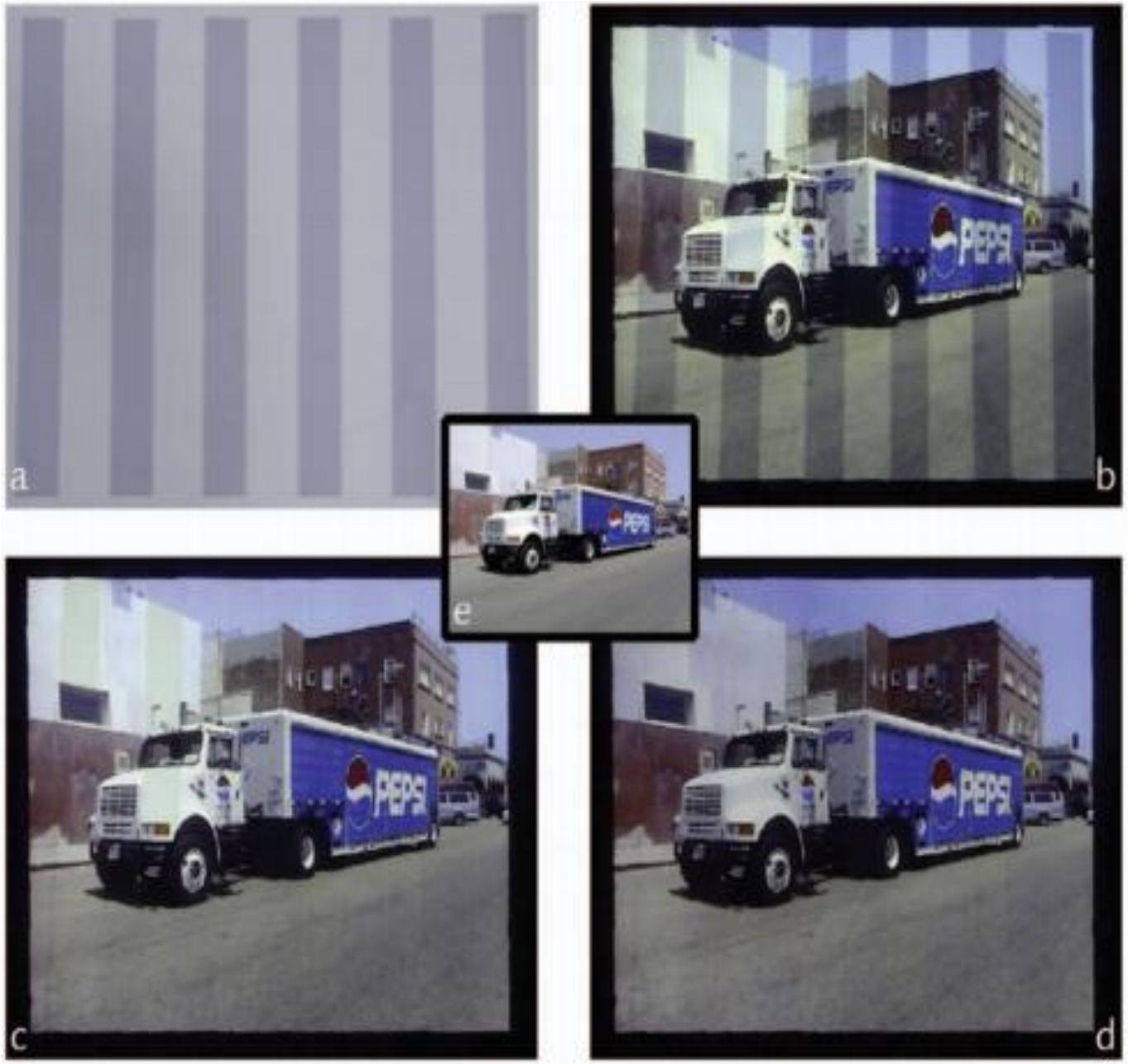
Motion stabilization
[Azuma et al.]



Handle blur
[Tyoura et al.]



Handle occlusion
[Percheim et al.]



Handle surface variations
[Bimber et al.]

Other factors

- Color
- Contrast and Brightness
- Display type (Video see through, Glass, Mobile, …)
- Stereoscopy
- Field of view

Google Glass study [Yu and Kim]

Torsion, Dual focus and Fatigue

[movie]

Other: Who Delivers where? [Jo and Kim]

2D video

Photorealistic character in AR

Photorealistic character in VR

Character avatar in VR

Character avatar in AR

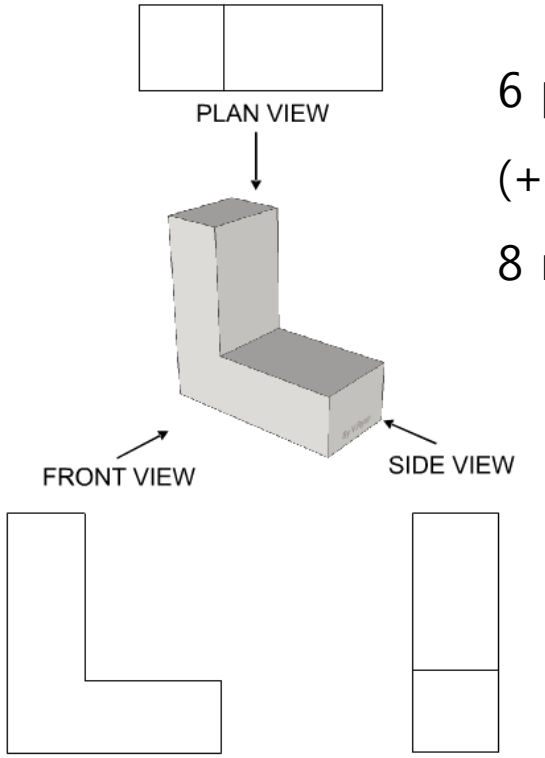
movie

Proposal for MAR standard – Occlusion issue

- Basic display requirement
  - ISO 9241 Series / Part 3 (Visual display requirements) addresses it
    - There may be some differences among various MAR displays (VST, OST (Glass), Mobile, Open Lens) but they might be similar to those between e.g. different desktop monitors, projectors and displays …
  - Likewise for performance specifications (clarity, legibility, stability, fatigue) under dynamic working conditions
    - Dynamic conditions once measured can be addressed by programmable and variable adjustments
- Truly MAR specific visualization specification? → Occlusion

Specify occlusion behavior – as part of *MAR content behavior*

- Augmentation occludes the target object
  - Eccentricity (where relative to the object)
  - Maximum coverage
  - Opacity
  - …
  - View angle dependent



6 principal directions

(+x –x +y –y+z –z) and

8 more diagonal directions

- Behavior::MARSGNode

Behavior objects specify the dynamics of the objects in the scene, e.g. time based or event based behaviour. Therefore, such behaviors can be expressed simply by relying on the usual script nodes. Behavior nodes, in addition to housing such flexible scripts, serve to abstract some typically used MAR “augmentation” behaviors (i.e. how virtual augmentation objects associated and spatially registered to a real physical object behave reacting to input and external events). The main purpose of such an abstraction is the ease of use, simplicity and thereby quick authoring. Behavior nodes are driven by external events/data produced from the MAR system components (such as the Sensor, Capturer, Tracker and Recognizer, see Clause 6.3) and accessible indirectly through the Mapping node (see Clause 7.4), and simulated by the MAR simulation engine (another important part of MAR system). The behaviour node will have associations with the content objects which are needed to drive and simulate the behaviour and those that are affected by the simulation as well.

Examples of often used augmentation behaviour are simple activating/deactivating of the augmentation to be visible, simple animation of them, highlighting effects, changing of transparency and color.

Behavior			
Access type	Data type	Attribute/Method name	Explanation
private	int	type	specifies the specific type of the behavior (see subclauses)
private	bool	isTriggered	Enabling or disabling the triggering of this Behavior node (0 for disabling it , and 1 for enabling it)
private	Mapping[]	eventSource	list/array of source Mapping node(s) that delivers and filters the event/data to this node from the MAR System components
private	MARSGNode[]	sourceObjects	list/array of source objects referred and used for behavior definition
private	MARSGNode[]	targetObjects	List/array of MArSGNodes that are affected by this Behavior node
private	VirtualEvent*[]	virtualTriggers	list/array of virtual events produced by the associated Mapping node for tri ggering this Behavior node
private	RealEvent*[]	realTriggers	list/array of real world physical events produced by the associated Mapping node for triggering this Behavior node.
private	string*	scriptFilePath	file path to a script if any.
public	Behavior*	Behavior()	constructor
protected	void	init()	initialization
private	void	triggerBehavior()	trigger the behaviour as specified in this node by updating and retrieving t he most recent events/data from the associated Mapping nodes and updat e the MAR scene structure
public	string	getScriptFilePath()	return the script file path
public	void	setScriptFilePath(string path)	set the script file path

Scenario (Marker or Image patch based, Video see-through)

Assumed R/V Root TGs are initialized by system automatically

We consider only MAR content side and do not consider anything about system side

Assumed R. Root

Assumed V. Root

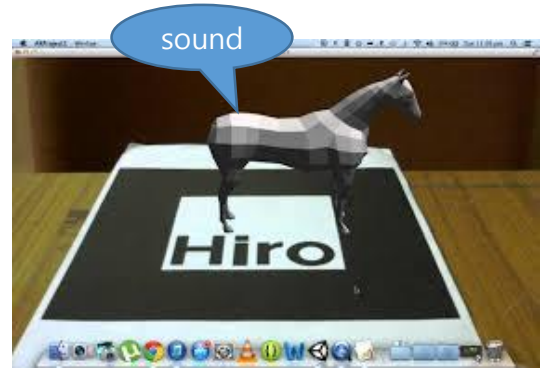
Marker RO

RTG

R

VTG

Viewpoint VO



R

Marker Tracker

Live BG VO

VTG

M

Horse VO

Sound VO

Live Video Capturer

System side doesn't know anything about MAR content

Front
- Location: Center
- Max coverage: 30%
- Opacity: 75%

Left side
- Location: Top left
- Max coverage: 40%
- Opacity: 10%

...

Behavior Visibility/ Sound

M

Touch Sensor

Touch Event

Camera Sensor

Sensor just senses some real world data