**ACM Multimedia Systems 2017, Taipei, Taiwan | June 20-23, 2017**


# Time Evaluation for the Integration of a Gestural Interactive Application with a Distributed Mulsemedia Platform

**SALEME, Estêvão Bissoli**, CELESTRINI, J. R., and SANTOS, C. A. S.
Federal University of Espírito Santo, Brazil

# Outline

- Introduction and motivation
- Experimental environment
  - Distributed mulsemedia platform
  - Gestural Interactive application
  - Test setup, design (procedure)
- Results
- Discussion and conclusion

# Intro and Motivation

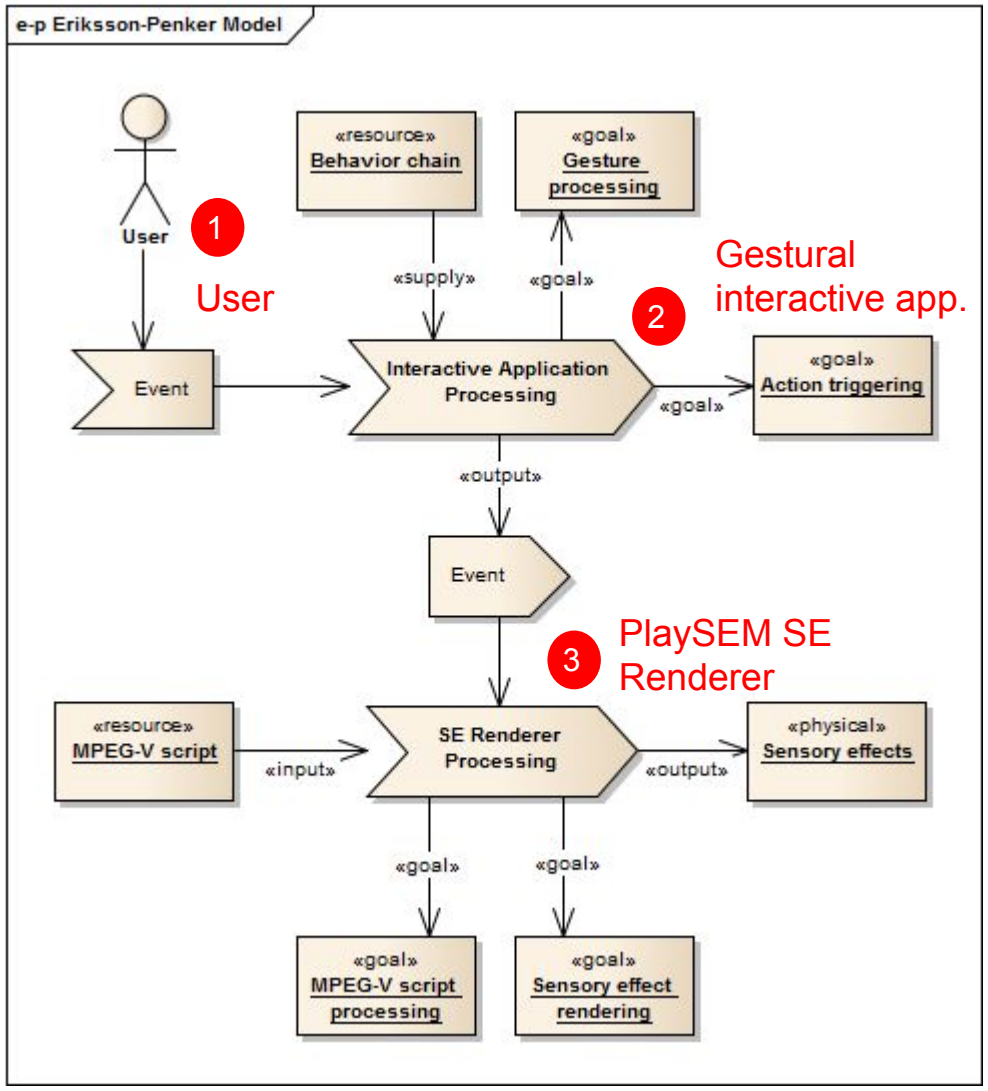
- Synchronization of **M**ul**S**e**M**edia (**M**ulti-**S**ensory **M**edia) has to do with transmission, production, and presentation of multiple signals (EG et al., 2015)
- For some sensory media, such as olfaction, the delay is acceptable for a few seconds (GHINEA et al., 2010; MURRAY et. al, 2013; MURRAY et. al, 2016) whereas others, such as visual and haptic, it can be stricter (MILLER, 1968; CARD et al., 1983; NIELSEN, 1993; YUAN et al., 2015)
- PlaySEM (distributed mulsemedia platform) -> has nearly no synchronization loss with timeline applications (SALEME et al., 2015)
- Delay is unknown when working with other kinds of applications (event-based) so far

# Our research

- We integrated the PlaySEM platform with a gestural interactive application presented by Santos et al. (2015), which is a simulation of a theatrical play where the presenter interacts through gestures to perform actions or sensory effects on the stage

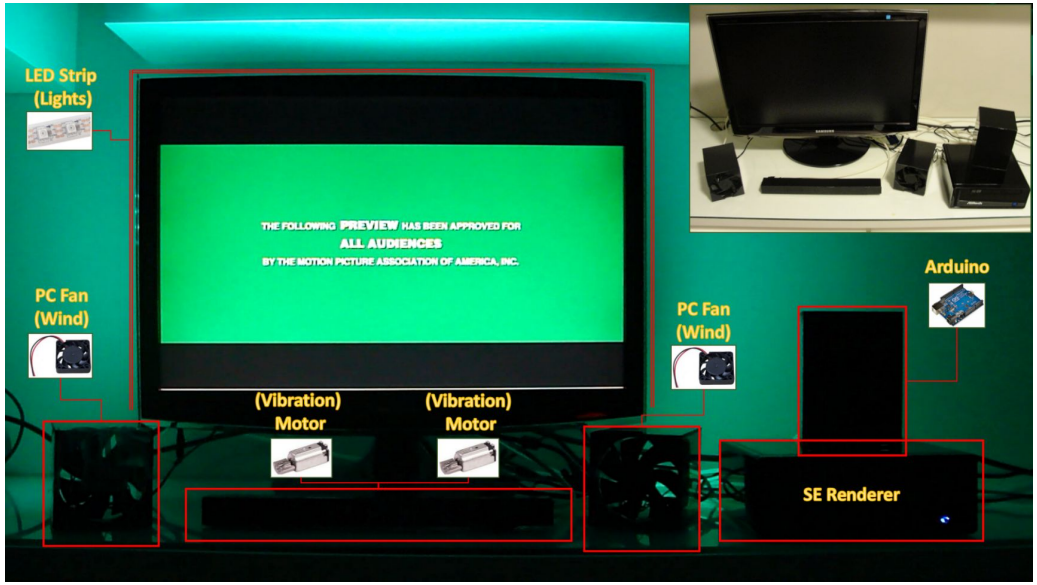
  - How long does the computational time (for transmission and production) of this integration take considering the distributed architecture?
  - What happens before the interactive application calls the sensory effects? What happens after that?
  - What can we do to improve it when necessary?

# Our scenario



# PlaySEM

- The Play**SEM** (**S**ensory **E**ffects **M**etadata) platform consists of a set of open source tools (based on SeSim) written in Java for playing and rendering sensory effects in the user environment
- 3 main components: SE Video Player, SE Renderer, and a microcontroller



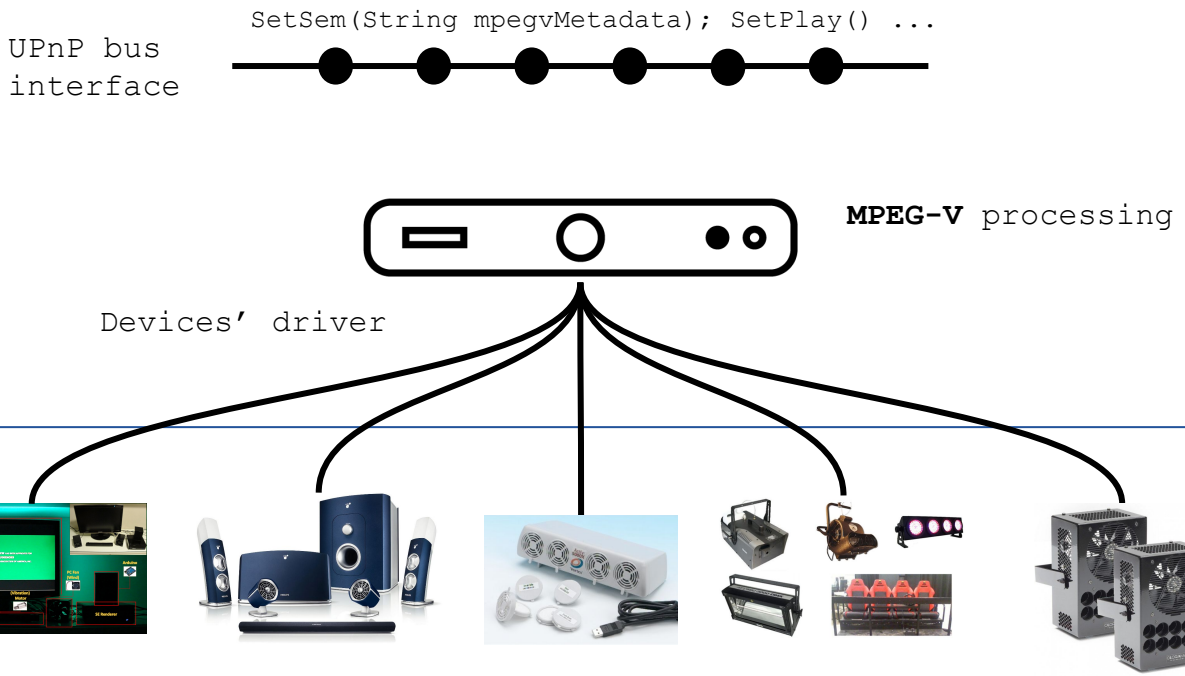

(SALEME et al., 2015)

# PlaySEM's architecture

**Applications**

SE Player, interactive apps, VR/AR apps, SEMP, Sensible Media Sim, VLC, MIT Sensory Fiction, etc

**PlaySEM SE Renderer**




**Devices**



PlaySEM hardware, Philips amBX, Vortex, CAVE System, Olorama Aromatizing, etc

# Interactive application

- Interactive spaces ⇒ well-defined physical environment composed of heterogeneous devices, applications, and users
- Santos et al. (2015) developed an interactive application in which a user interacted through gestures with a system to control the scenes of a theatrical play producing lighting, wind and vibration effects
- Other examples of interactive environments:

**MEDIATE, multisensorial interactive environment for children with autism**
Source: http://www.annacarreras.com/eng/mediate/

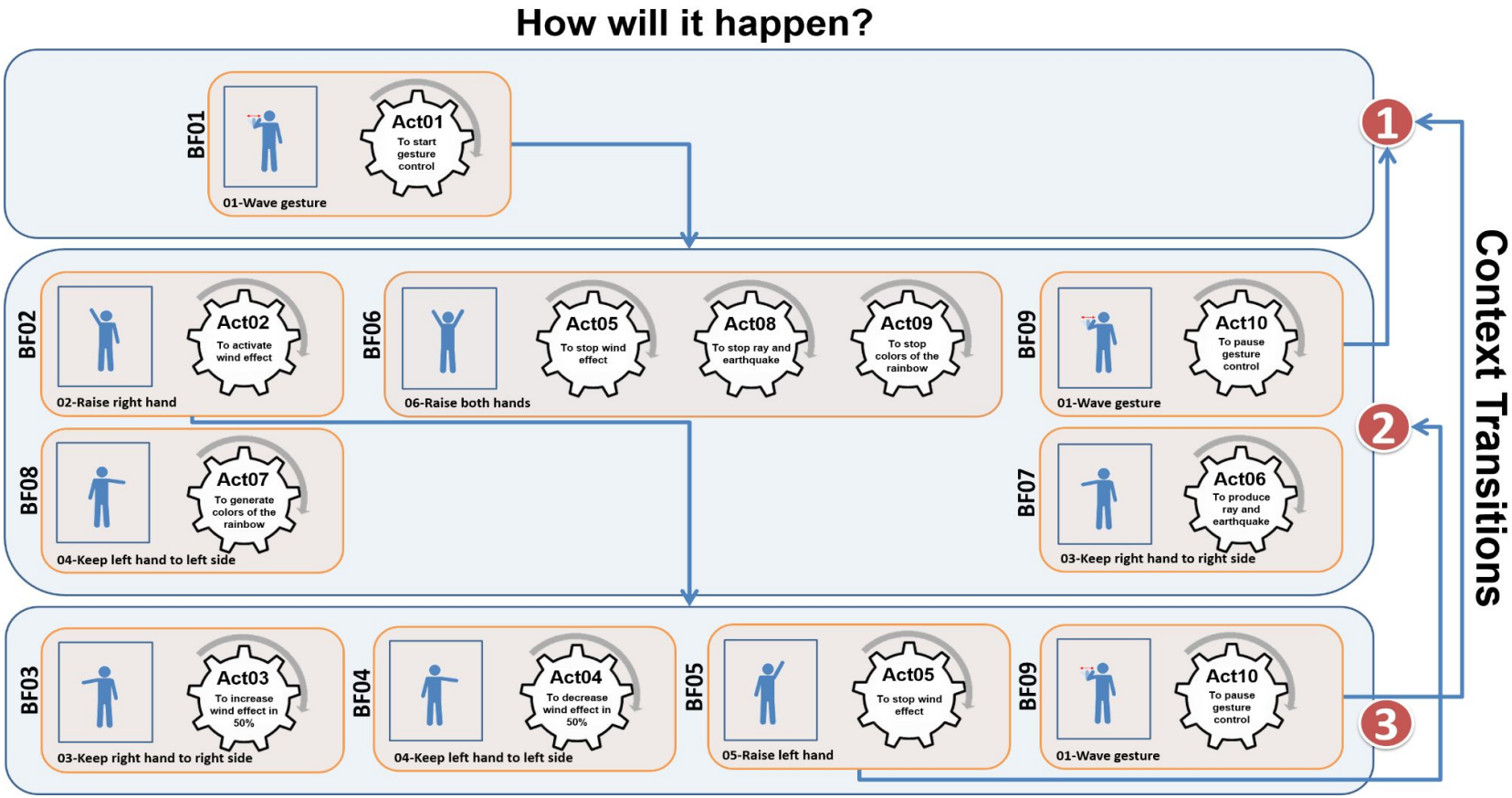
**The Exhale pavilion**
Source: https://www.e-architect.co.uk/miami/exhale-pavilion

# Interactive application

- 3W1H (Where, What, Why, How) approach for developing interactive applications (REHEM NETO et al., 2015)
- Gap: ad-hoc and unstructured solutions
- It formalizes a set of stages for developing interactive scenarios
- Based on the principle of stimuli-response
- It combines actions that change the scenario with specific events are recognized

# Interactive application

- Gestures, actions and their relationship




(SANTOS et al., 2015)

# Interactive application

**Behavior Chain (Rules) as an XML file**

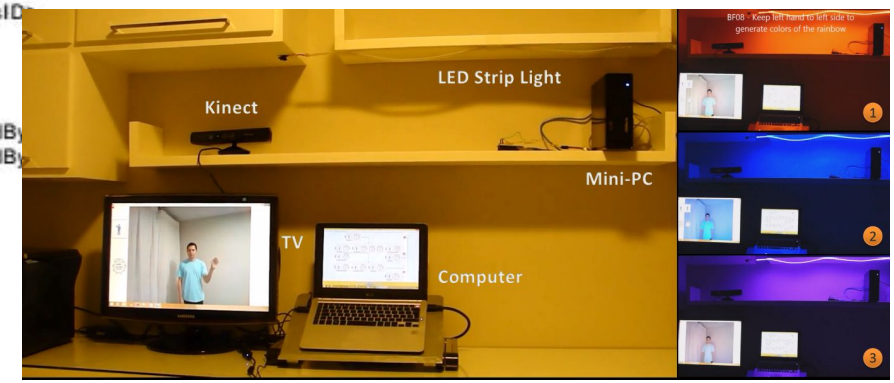
```
<?xml version="1.0" encoding="UTF-8"?>
<Behavior (...)
    <interaction ID="INT01">
        <name>Wave gesture</name>
    </interaction> (...)
        <interaction ID="INT06"> (...)
        <action ID="ACT01">
        <name>Starts gesture control</name>
    </action> (...)
        <action ID="ACT10"> (...)
    <behaviorFrame ID="BF01">
        <name>BF01</name>
        <interactionID>INT01</interactionID>
        <actionID>ACT01</actionID>
    </behaviorFrame> (...)
    <behaviorFrame ID="BF09"> (...)
    <expectancy ID="exp01">
        <name>Gesture control deactivated</name>
        <behaviorFramesID>BF01</behaviorFramesID>
    </expectancy>
    <expectancy ID="exp02"> (...)
        <behaviorFramesID>BF02 BF06 BF07 BF08 BF09</behaviorFramesID
    </expectancy>
    <expectancy ID="exp03"> (...)
    <behaviorChain ID="BC01">
        <expectanciesID>exp01 exp02 exp03</expectanciesID>
        <expectancyTransition source="exp01" target="exp02" causedB
        <expectancyTransition source="exp02" target="exp01" causedB
    </behaviorChain>
</Behavior>
```

(SANTOS et al., 2015)

**Excerpt of an action (MPEG-V script)**

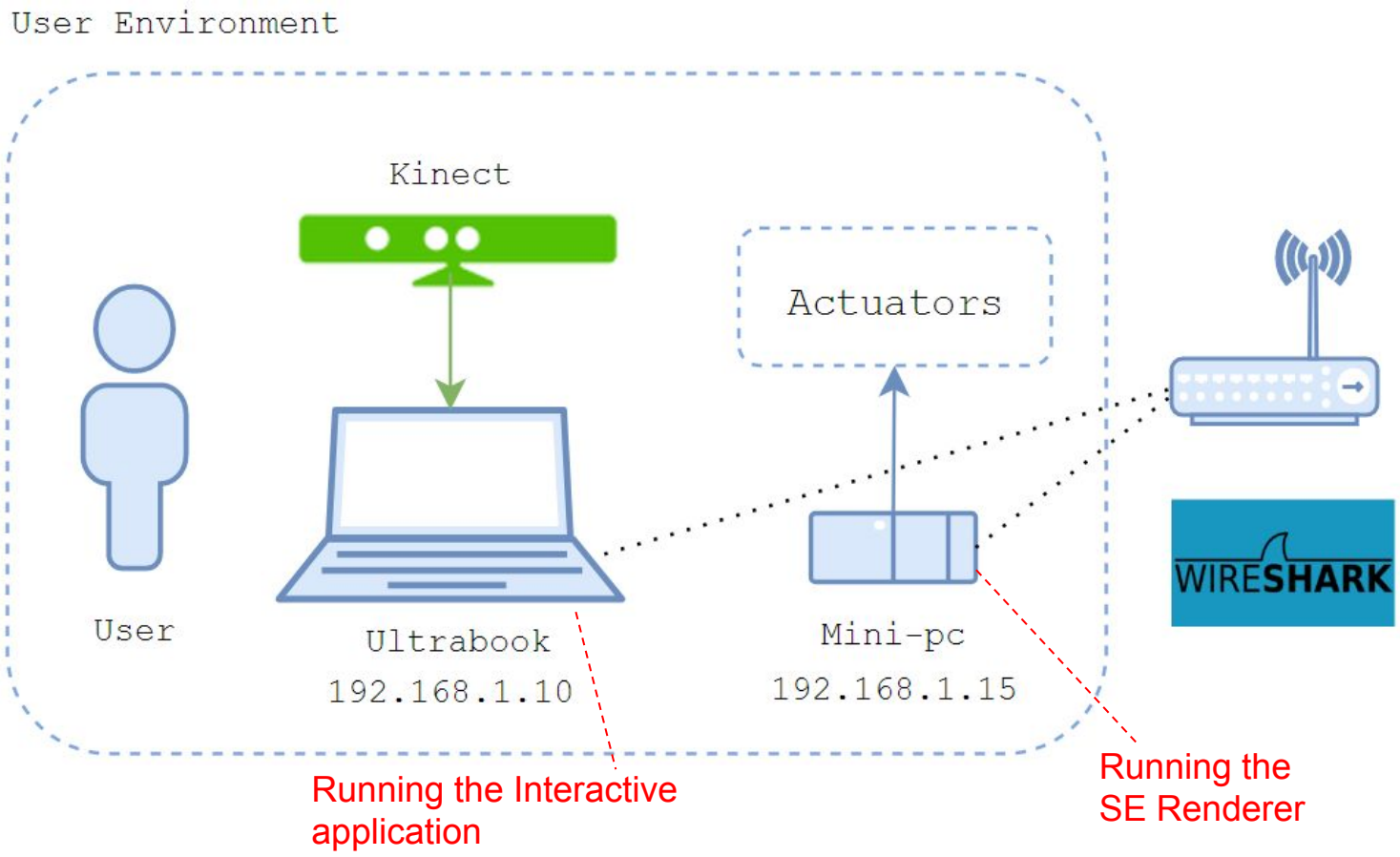
```
<ns2:Effect xsi:type="ns3:LightType" ... si:pts="1000" ns3:color="urn:mpeg:mpeg-v
        :01-SI-ColorCS-NS:red"/>
<ns2:Effect xsi:type="ns3:LightType" ... si:pts="99000" ns3:color="urn:mpeg:mpeg-v
        :01-SI-ColorCS-NS:international_orange"/>
<ns2:Effect xsi:type="ns3:LightType" ... si:pts="198000" ns3:color="urn:mpeg:mpeg-v
        :01-SI-ColorCS-NS:amber"/>
<ns2:Effect xsi:type="ns3:LightType" ... si:pts="306000" ns3:color="urn:mpeg:mpeg-v
        :01-SI-ColorCS-NS:dark_pastel_green"/>
<ns2:Effect xsi:type="ns3:LightType" ... si:pts="432000" ns3:color="urn:mpeg:mpeg-v
        :01-SI-ColorCS-NS:blue"/>
<ns2:Effect xsi:type="ns3:LightType" ... si:pts="576000" ns3:color="urn:mpeg:mpeg-v
        :01-SI-ColorCS-NS:han_purple"/>
<ns2:Effect xsi:type="ns3:LightType" ... si:pts="729000" ns3:color="urn:mpeg:mpeg-v
        :01-SI-ColorCS-NS:violet_ryb"/>
<ns2:Effect xsi:type="ns3:LightType" activate="false" si:pts="900000"/>
```

**Environment**

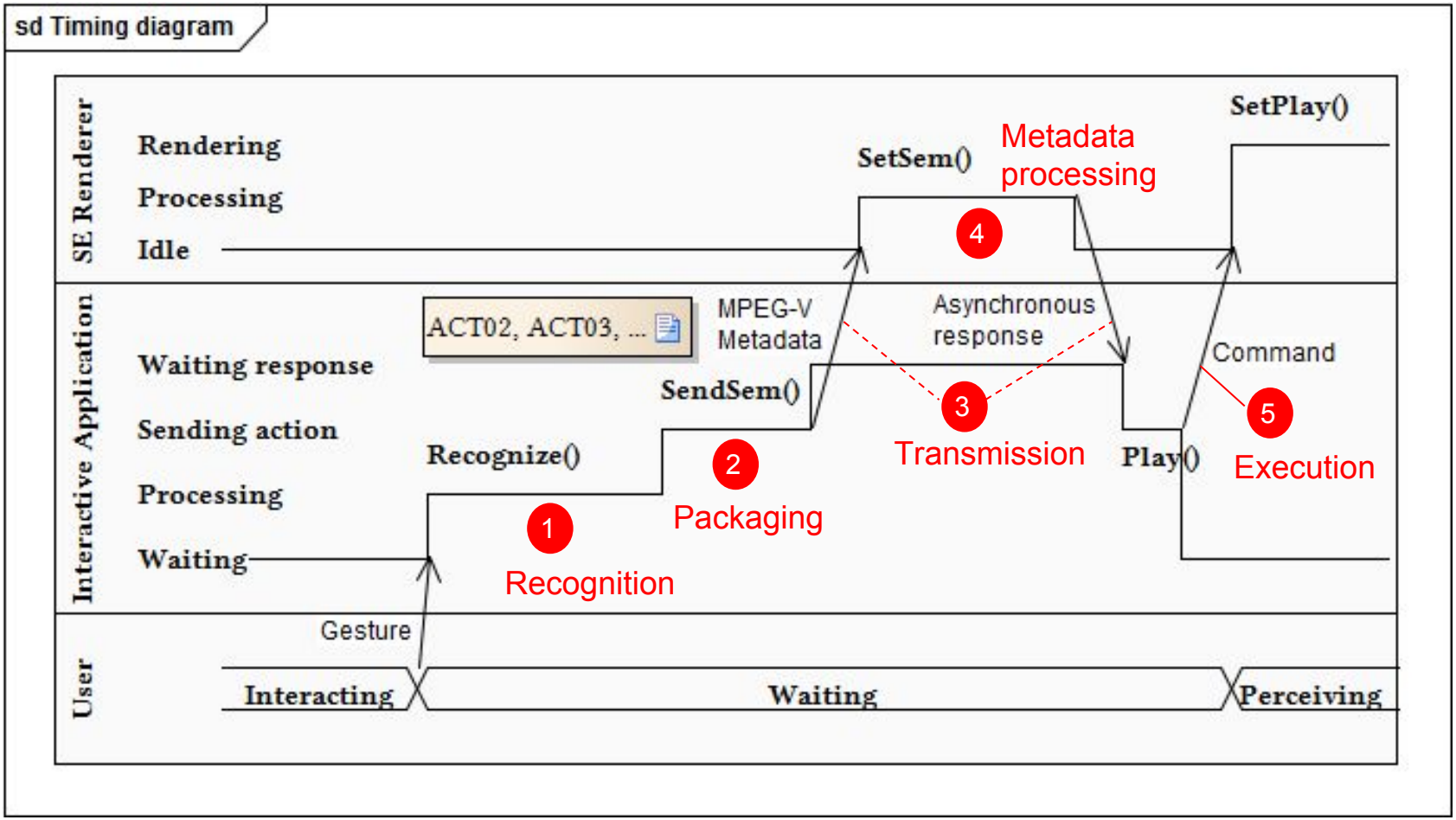


# Test setup

- Arrangement of devices, tools and network



# Experimental Procedure



# Results

| Interaction | | | Average time | |
|---|---|---|---|---|
| Behavior Frame | Gesture | Actions | Recognize (ms) | SendSem (ms) |
| BF02 | Raise right hand | ACT02 | 9,39 | 13,21 |
| BF03 | Keep right hand to right side | ACT03 | 9,80 | 13,76 |
| BF04 | Keep left hand to left side | ACT04 | 9,76 | 13,33 |
| BF05 | Raise left hand | ACT05 | 9,47 | 12,73 |
| BF06 | Raise both hands | ACT05, ACT08, ACT09 | 10,45 | 27,70 |
| BF07 | Keep right hand to right side | ACT06 | 11,76 | 17,76 |
| BF08 | Keep left hand to left side | ACT07 | 10,19 | 15,28 |

1 Recognition

2 Packaging

- Before the sensory effects
  1. Gesture recognition takes around 10ms (processing)
  2. Packging MPEG-V scripts takes from 13 to 28ms (processing)
     - Increase in delay when having more actions - UPnP request
     - Send all scripts as soon as possible and afterward just call for action(s) -> it would remove packaging time (it requires changes in the SE Renderer)

# Results

1 Transmission

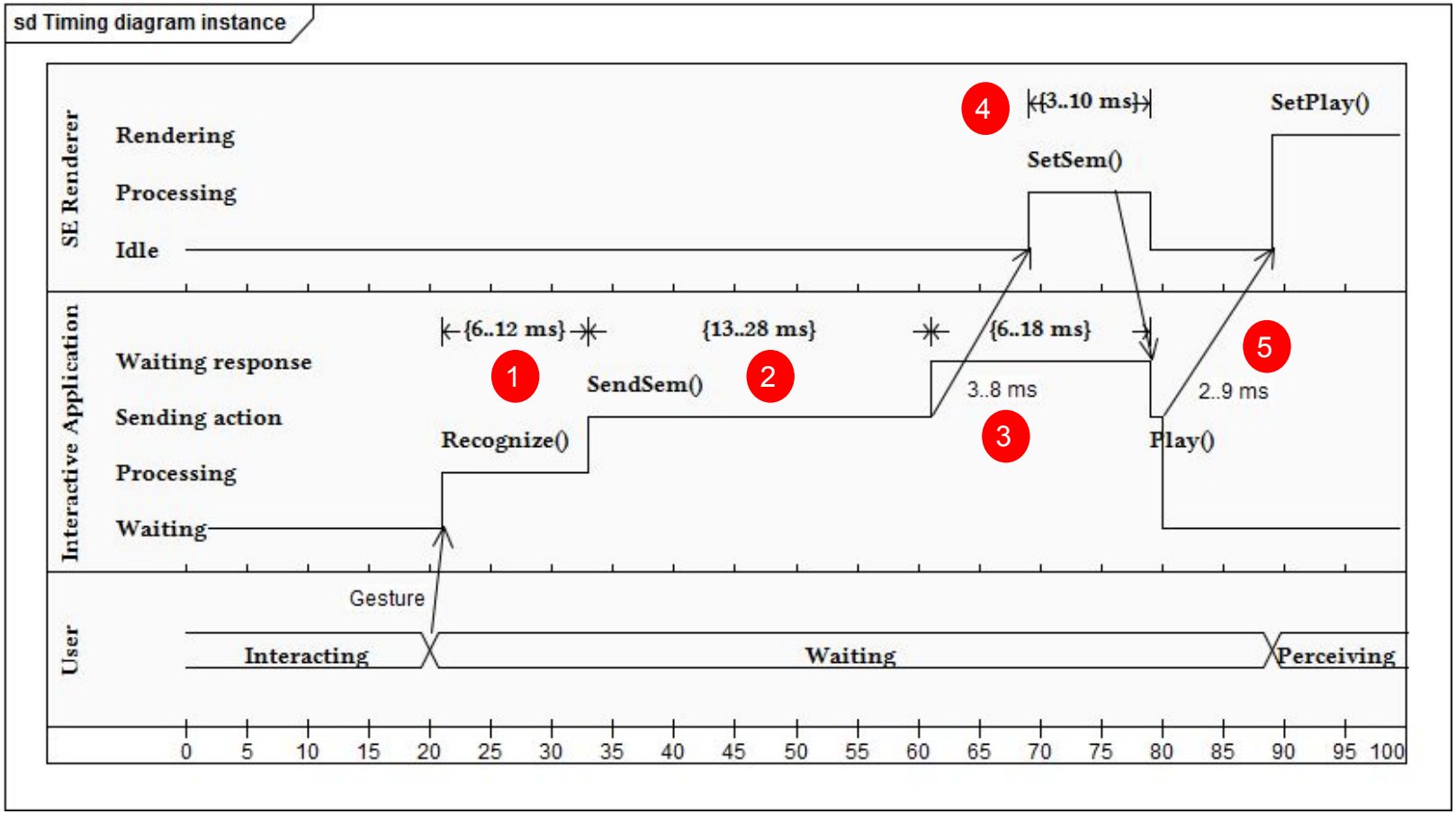
Metadata processing 2

| | Wired network transmission | | Wireless network transmission | | Processing |
|---|---|---|---|---|---|
| Action | SetSem (ms) | SetPlay (ms) | SetSem (ms) | SetPlay (ms) | SetSem (ms) |
| ACT02 | 7,84 | 9,22 | 34,51 | 18,27 | 9,58 |
| ACT03 | 5,49 | 6,43 | 39,27 | 13,15 | 8,04 |
| ACT04 | 4,86 | 5,18 | 33,00 | 8,23 | 5,31 |
| ACT05 | 5,51 | 3,77 | 33,48 | 6,28 | 7,44 |
| ACT06 | 4,88 | 3,84 | 35,85 | 8,81 | 6,24 |
| ACT07 | 3,64 | 3,78 | 37,46 | 4,98 | 5,10 |
| ACT08 | 2,88 | 1,90 | 35,51 | 7,86 | 3,12 |
| ACT09 | 4,91 | 3,05 | 32,74 | 5,32 | 3,77 |

- Transmitting and processing
  1. Transmission time is affected by the type of network as expected
     - SetSem conveys data; SetPlay is just a command
     - Send all scripts after handshaking -> it would suppress the command SetSem for transmission as well as for processing
  2. Processing is not expensive; it reaches up to 10ms
     - It could be optimized replacing the programming language

# Results

- Range of average time (wired network) = 27ms to 67ms



# Discussion and Conclusion

- The range of 27ms to 67ms reveals the average time taken for dealing with computational aspects before presenting an effect.
- Previous time + devices' activation time + perceptual time must be considered when developing a distributed mulsemedia solution
- Aspects to be observed for improving global response time:
  - Avoid sending MPEG-V metadata to the platform in real-time
  - Reduce the number of messages exchanged on the network
  - Consider using a wired network (test setup)
  - Use a more efficient programming language to convert MPEG-V metadata into commands for sensory effect devices
- Future work: an infrastructure for mulsemedia solutions concerning integration processes, software, and hardware inspired by CPS

# References


Ragnhild Eg, and Carsten Griwodz. 2015. **Approaching New Limits of Synchrony with Multi-Sensorial Media**. IEEE MMTC R-Letter 7-9, 10/2015, IEEE

E. B. Saleme and C. A. S. Santos. 2015. **PlaySEM: A Platform for Rendering MulSeMedia Compatible with MPEG-V.** In Proceedings of the 21st Brazilian Symposium on Multimedia and the Web (WebMedia '15). ACM, New York, NY, USA, 145–148. https://doi.org/10.1145/2820426.2820450

N. Murray, B. Lee, Y. Qiao, and G. M. Muntean. 2016. **The influence of human factors on olfaction based mulsemedia quality of experience.** In 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX). 1–6. https://doi.org/10.1109/QoMEX.2016.7498975

G. Ghinea and O. A. Ademoye. 2010. **Perceived Synchronization of Olfactory Multimedia**. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 40, 4 (July 2010), 657–663.

C. A. S. Santos, A. N. R. Neto, and E. B. Saleme. 2015. **An Event Driven Approach for Integrating Multi-sensory Effects to Interactive Environments**. In Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on. 981–986. https://doi.org/10.1109/SMC.2015.178

R. B. Miller. 1968. **Response Time in Man-computer Conversational Transactions**. In Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I (AFIPS '68 (Fall, part I)). ACM, New York, NY, USA, 267–277.

S. K. Card, A. Newell, and T. P. Moran. 1983. **The Psychology of Human-Computer Interaction**. L. Erlbaum Associates Inc., Hillsdale, NJ, USA

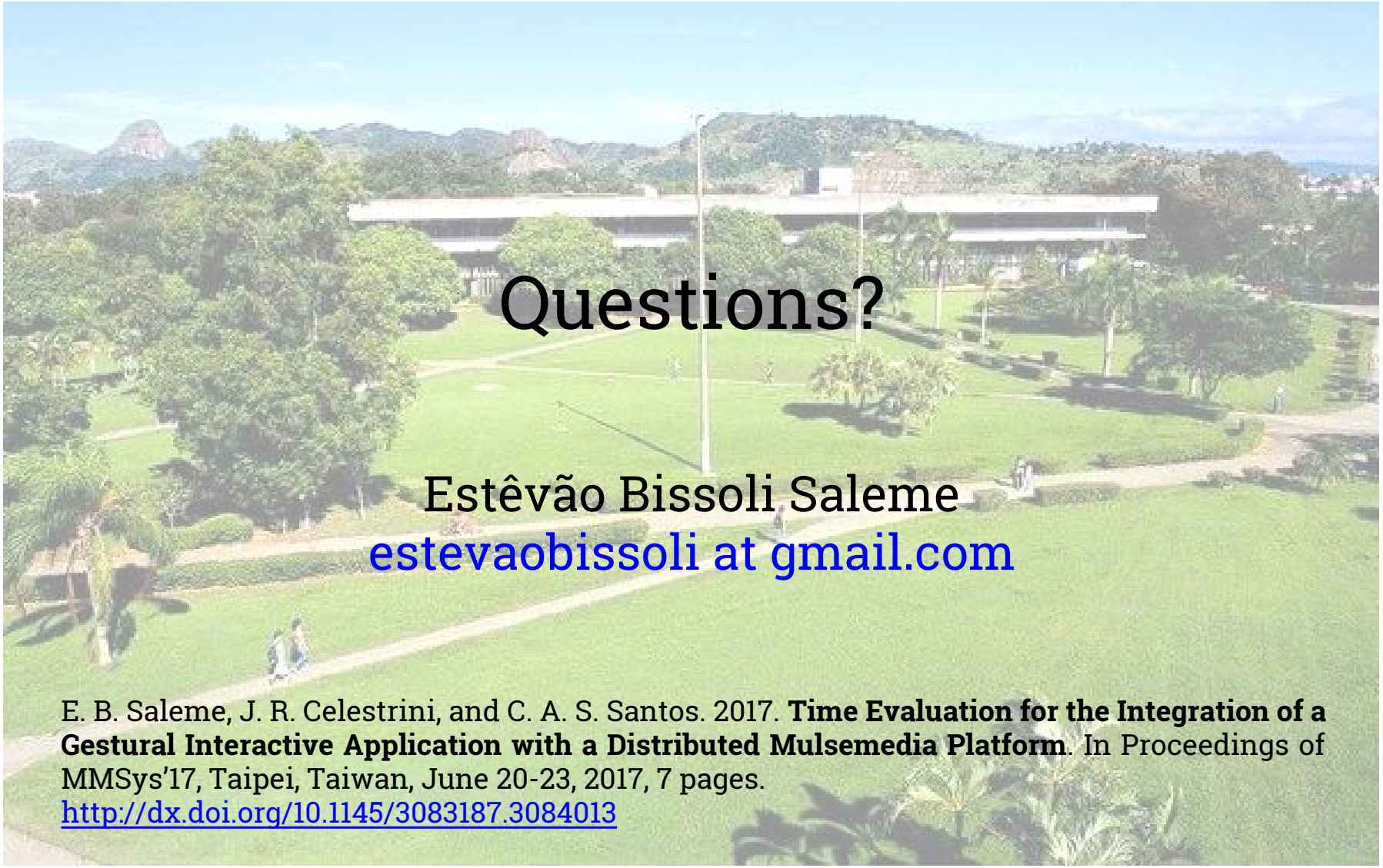
J. Nielsen. 1993. **Usability Engineering**. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA

Z. Yuan, T. Bi, G. M. Muntean, and G. Ghinea. 2015. **Perceived Synchronization of Mulsemedia Services.** Multimedia, IEEE Transactions on PP, 99 (2015), 1–1. https://doi.org/10.1109/TMM.2015.2431915

Niall Murray, Yuansong Qiao, Brian Lee, A. K. Karunakar, and Gabriel-Miro Muntean. 2013. **Subjective evaluation of olfactory and visual media synchronization**. In Proceedings of the 4th ACM Multimedia Systems Conference (MMSys '13). ACM, New York, NY, USA, 162-171. DOI=http://dx.doi.org/10.1145/2483977.2483999

A. N. Rehem Neto ; C. A. S. Santos; E. B. Saleme; J. S. S. Menezes . **Where, What, Why and How - 3W1H: Uma abordagem prática para desenvolvimento de ambientes interativos.** In: XIV Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais 2015, Salvador, Brazil. IHC. 2015

# Thanks for your attention!




E. B. Saleme, J. R. Celestrini, and C. A. S. Santos. 2017. **Time Evaluation for the Integration of a Gestural Interactive Application with a Distributed Mulsemedia Platform**. In Proceedings of MMSys'17, Taipei, Taiwan, June 20-23, 2017, 7 pages.
http://dx.doi.org/10.1145/3083187.3084013