

# Device characterization for conditional encoding

VQEG Madrid 2018

**Juan Casal<sup>†</sup>**, **Mateo Cámara<sup>\*</sup>**, **César Díaz<sup>\*</sup>**, **Jorge Ruano<sup>†</sup>** and **Narciso García<sup>\*</sup>**

[juan.casal\\_martin@nokia.com](mailto:juan.casal_martin@nokia.com)

[narciso@gti.ssr.upm.es](mailto:narciso@gti.ssr.upm.es)

<sup>†</sup>Nokia Spain  
Madrid

<sup>\*</sup>Grupo de Tratamiento de Imágenes (GTI)  
Universidad Politécnica de Madrid

# Presentation scheme

- Mobile video
- Nokia's conditional encoding
- Motivation for the subjective evaluation
- Content selection and spatial ladder
- Test environment and procedure
- Some results
- Conclusions
- Demo & discussion



# Mobile Video Evolution

A perfect storm

- ❑ Mobile Video demand grows and grows
- ❑ Unlimited Data Plans are the next commercial proposition
- ❑ Huge investments will be required to cope with the increasing demand
- ❑ More traffic does not mean more money. Revenues per data traffic will remain stable while consumption increases

Efficiencies in content encoding mean huge savings in network investments



Will you stand looking ?

# Mobile ABR video service optimization

Our target - Optimize mobile video delivery in real time

## Technical Challenges

- ❑ Content aware optimization of an ABR service in terms of quality-bitrate
  - ❑ References and existing ABR service
  - ❑ Maintains the quality ladder of the original layers
  - ❑ Optimization in terms of quality-bitrate by dynamically modifying resolution and bitrate.
  - ❑ Optimization done in a segment or sequence basis.
- ❑ Accelerate extraction of video complexity features. By applying IA techniques.

## Research Challenges

There is a need to differentiate client devices

## Reference Resolution

- ❑ What is the resolution we should target for a device depending of the size of the screen?

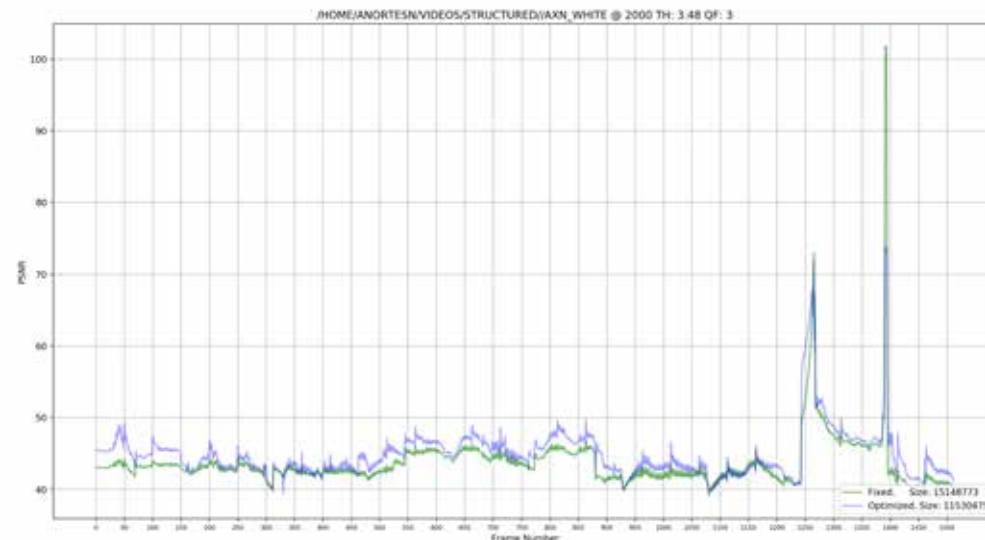


# Conditional Encoding Project

Real time content aware encoding for mobile devices

- ❑ CE is functionally thought as an intelligence layer on top of the ABR video encoders
  - ❑ Accelerates content aware encoding.
  - ❑ Extracts video complexity features in real time by using IA techniques.
  - ❑ Optimizes video services in terms of quality and bitrate, in real time.
  - ❑ Creates optimal configurations (resolution/bitrate), on a segment basis, for the ABR video encoders.

Objective results show above  
30% bitrate reduction  
maintaining the quality



# Motivation for the subjective evaluation

- Multimedia delivery based on Nokia conditional encoding.
- Need for device characterization
  - Usually: device multimedia interface
  - Our approach: user experience -> device rendering capabilities (display/loudspeaker)
  - Remember Full HD versus HD Ready (same interface, but different display)
- Initial activity on nomadic/mobile multimedia users
  - Mobile devices (smartphones or tablets)
  - Static (w.r.t. the platform) consumption (“nomadic use”)

## Content selection

- Looking for adequate temporal and spatial complexity.
- Pre-pilot choice of four representative sequences (Fish, Maldives, Goku, CHiPs)
  - Discarded due to ex-post evaluation
- New contents (UHD-4K resolution) for the study:
  - Low spatial and temporal complexity: Captain America (Gameplay)
  - Low temporal and high spatial complexity: Venice
  - High temporal and low spatial complexity: Skate
  - High temporal and spatial complexity: Football
  - Medium temporal and spatial complexity: India

# Spatial resolution ladder

- Sources:
  - Apple Technical Note 224
  - Netflix Per-Title Encode Optimization
- Decision: keep 16:9 frame aspect ratio along the spatial ladder
- Five-step spatial resolution ladder
  - 1920x1080p (Full HD)
  - 1280x720p (HD ready)
  - 960x540p (Wide SD, approx.)
  - 640x360p (Wide HVGA, approx.)
  - 480x270p (Wide QVGA, approx.)
- Content processing: downscaling - compression - decompression - upscaling - storage



# Devices

- LG-P720
  - Screen: 4.3 inches with 800x480 pixel resolution
- Samsung A3 (2017)
  - Screen: 4.7 inches with 1280x720 pixel resolution
- Samsung Galaxy S7
  - Screen: 5.1 inches with 2560x1440 pixel resolution
- iPad Air 2
  - Screen 9.7 inches with 2048x1536 pixel resolution (4:3 ratio)
  - Use of a letterbox display approach to show 16:9 content
  - So, equivalent to 2048x1152 within 8.9 inches



# Test considerations

- Classical audiovisual consumption and, therefore, QoE assessment:
  - Seated subjects looking at a fixed screen from a pre-set distance
  - Content narrative governing subject attention
  - Audio (very) important for subject opinion score
- Presuming audiovisual engagement, therefore looking for opinion scores



## Test environment

- Individual multimedia consumption
  - Audiovisual (video plus audio) content
  - Use of earplugs → Acoustic isolation from the environment
- Seated assessment and sequential evaluation on different devices
- User to display distance decided by user preferences
  - Mainly: device on the table
  - Some cases: device hand-held
  - So, around 40 cm
- Fifty observers (30 male, 20 female, average 22 years old)
- Visual acuity pre-screening and opinion scores data post-screening (four tests dropped)

## Environment and device settings

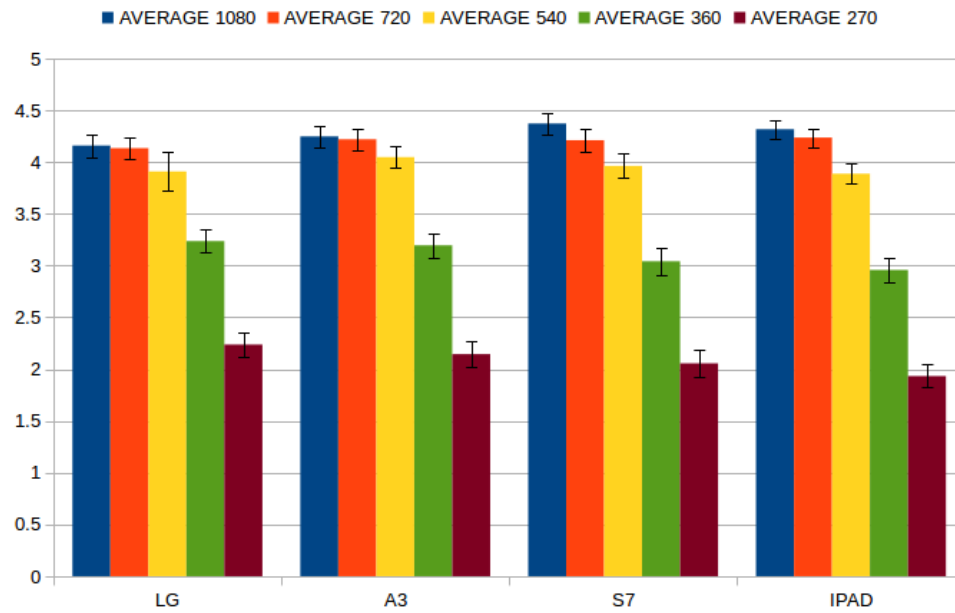
- Meeting room with neutral background
- Room illumination measures:
  - Front background ~100 lux
  - Left/right background (door) ~50 lux
  - Right/left background (window) ~250 lux (sunny day before blind attenuation) ~100 lux (cloudy day)
- Device illumination settings:
  - Brightness at 90% of maximum for smartphones
  - Content brightness on smartphones:
    - Grey background ~110 lux, Adaptation content (Sloth) ~70 lux
    - C.America ~170 lux, Venice ~70 lux, Skate~120 lux, Football~180 lux, India ~70 lux
  - Automatic brightness for iPad

## Test procedure

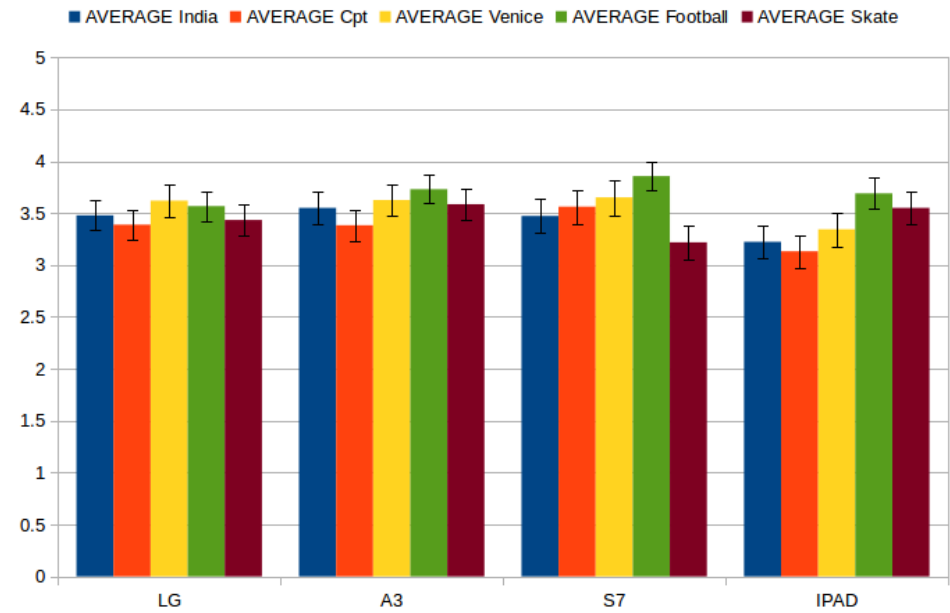
- Unrelated with the evaluation activity, ... although large audiovisual experience
- Explanation of the assessment procedure (ACR rating)
- For each device:
  - Initial (adaptation) neutral sequence with smooth music
  - User adjustment of earplugs sound level to his/her comfort level
  - Presentation of best and worst qualities
  - Random sequence of contents and qualities of ten seconds
  - Four seconds grey level sequence between tests sequences for voting time
- Final debriefing asking for comments and suggestions
- Overall assessment time about 30 minutes
- Test conducted in association with another test on variable player reproduction speed  
No pre-set order of the two assessments

# First results of the study (MOS)

Average regarding quality

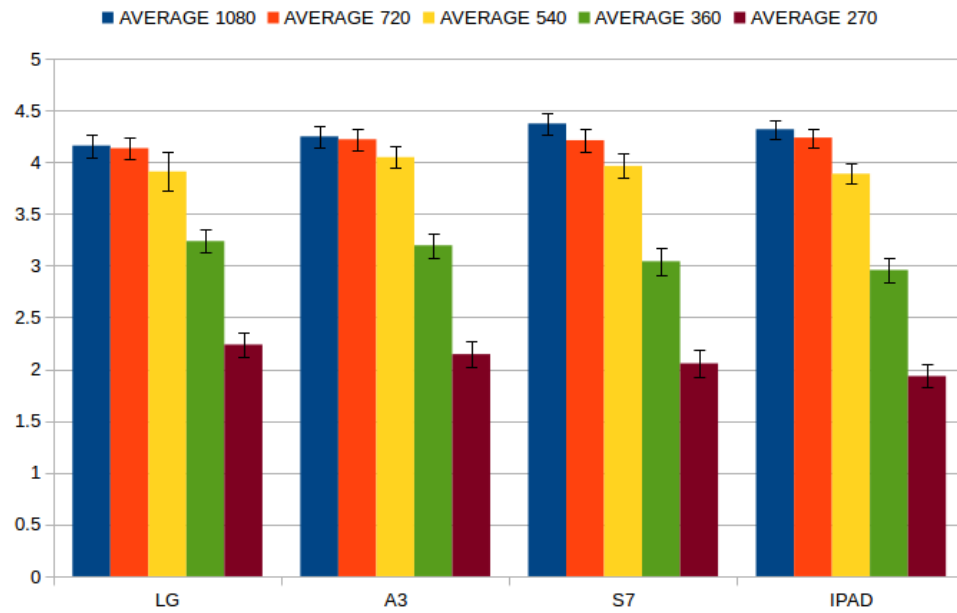


Average regarding content

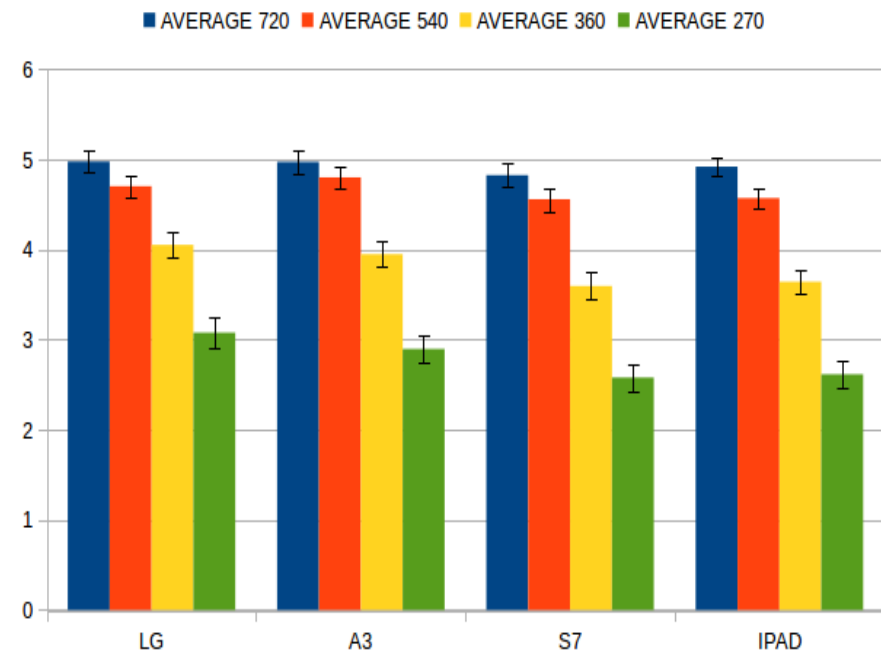


# First results of the study (MOS/DMOS)

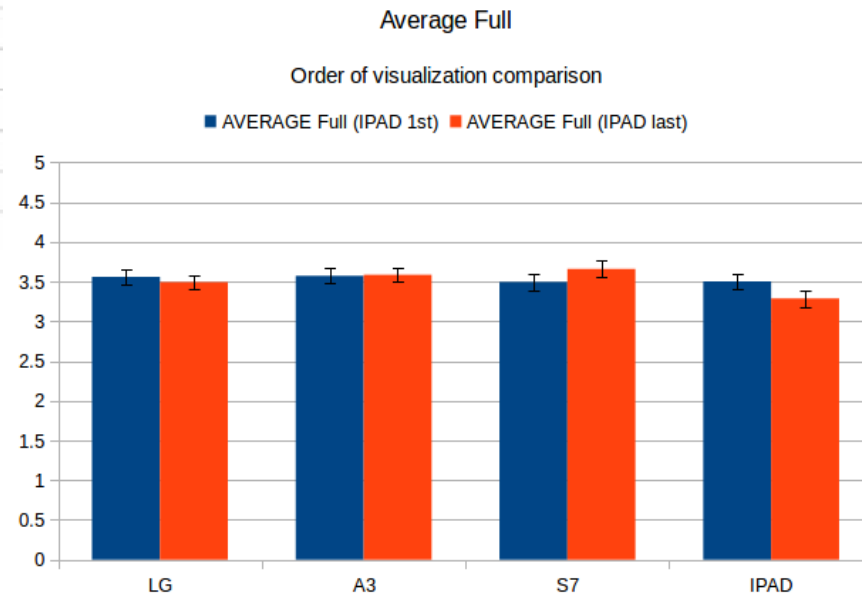
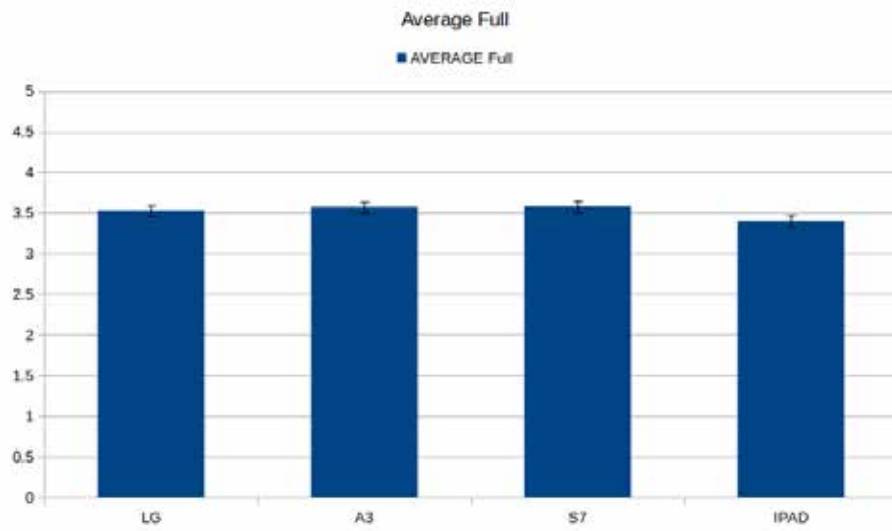
Average regarding quality



DMOS Average regarding quality

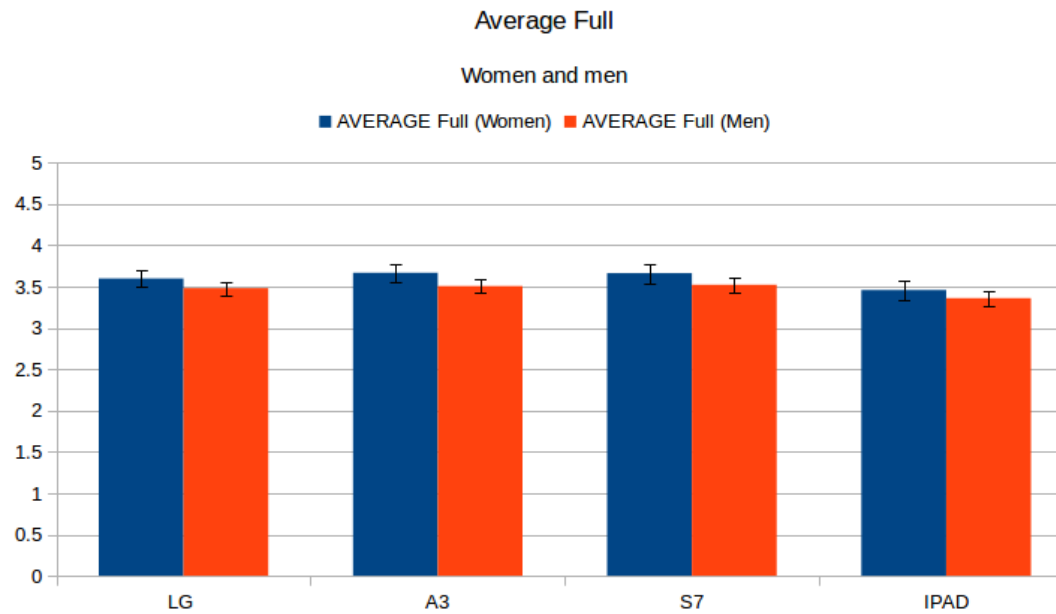


# First results of the study





# Curious initial result !?!?



## Main debriefing comments

- Satisfaction with the display of “best” and “worst” before the beginning of the test.
  - Help to establish a grading scale, mainly in previously not-used devices
- Difficulties in the evaluation of low spatial complexity contents.
  - Captain America and Skate
- Device order is important, mainly between smartphones and tablets.
- Possible biases due to repetition and personal preferences on device characteristics.
- Influence of content selection.

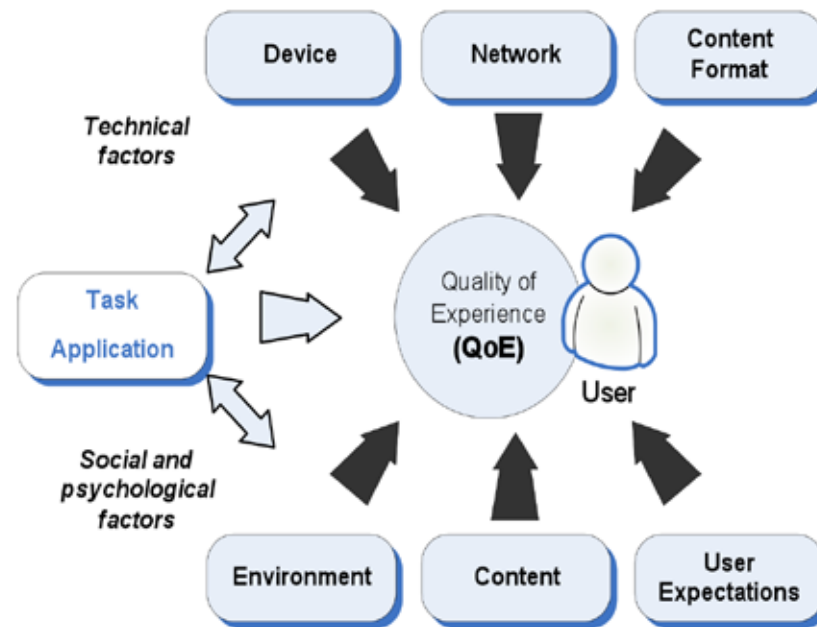
## Conclusions and future work

- Identification of the maximum perceptible resolution in different devices
- Statistical analysis on observers subsets:
  - First assessment versus second assessment (possible tiredness and/or naïve-less)
  - First half versus second half (possible training drift due to trainer boredom)
  - Random split into two subsets (experiment consistency)
  - Influence of the device order (memory effect?)
  - ... other unexpected results
- Minimum number of observers for reliable conclusions

## CE Conclusions and future work

- Subjective Validation of CE Mobile.
- Sequence based optimization. Will we get further improvements?
- Improve 4K scenario, creating 4K-like channels.

Finally, ... do not forget!



# Nokia's Conditional Encoding Demo

## Questions – Discussion – Debate - ...

Grupo de Tratamiento de Imágenes (GTI)  
Universidad Politécnica de Madrid

Nokia Spain  
Madrid

This work has been partially supported by the Ministerio de Economía, Industria y Competitividad (AEI/FEDER) of the Spanish Government under project TEC2016-75981 (IVME) and the Spanish Administration agency CDTI under project IDI-20170572 (BUSTOP)