

Aligning subjective tests using a low cost common set

Y. Pitrey, U. Engelke, M. Barkowsky,
R. Pepion, P. Le Callet

LUNAM University of Nantes
IRCCyN Image & VideoCommunication Group

Motivation

Corpus effect

Each experiment is a unique environment.

No direct comparison possible.

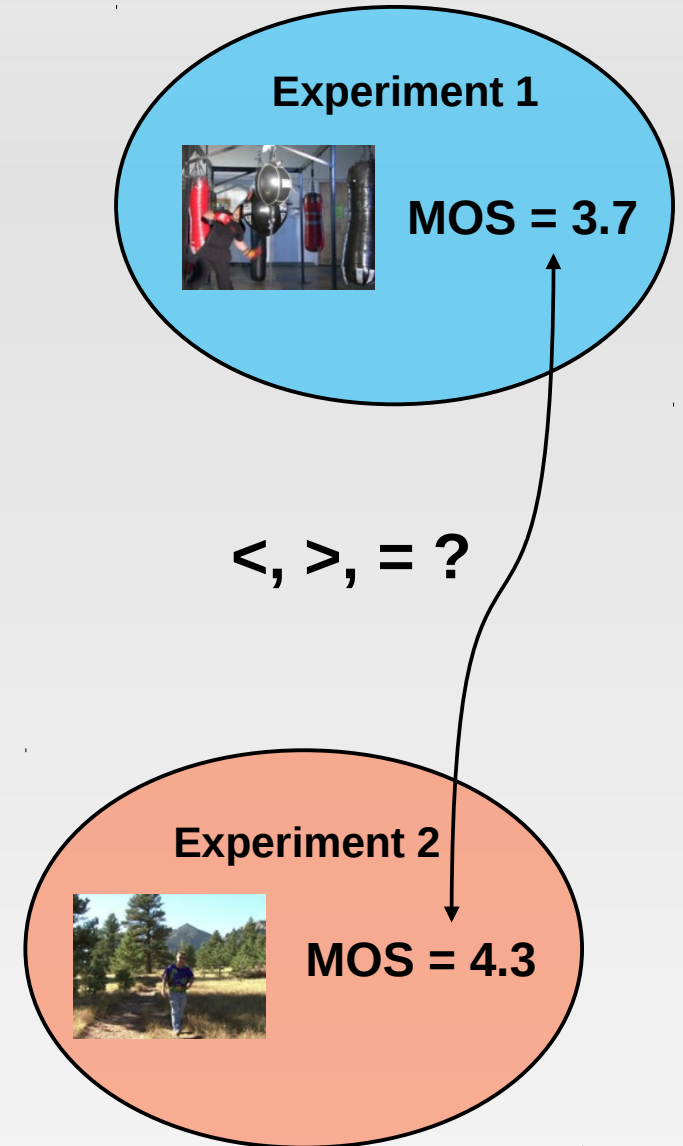
Common set of video sequences

Anchor conditions

Applications

Compare experiments

Merge experiments



Align experiments : a low-cost recipe

1. Ingredients : 3 subjective experiments
2. Design of the common set
3. Choosing a reference experiment
4. Mapping onto the reference experiment
5. Comparing experiments

Ingredients

3 subjective experiments on SVC transmission

2009 → 2011
IRCCyN (SVC4QoE project)

Global configuration :

2 SVC layers (VGA@30fps)
Part of Highest layer is lost
Upscale Base layer for error-concealment

9 video contents
ACR 5-level scale
Standard viewing conditions

TIMELINE



T1

SVC Error-concealment

Constant bitrate
1-second network impairment
3 error-concealment methods
2 fps for base layer

T2

SVC coding artifacts

Constant QP : 16 HRCs
Base layer upscale : 4 HRCs

T3

imprmt distribution

Constant QP HRCs
Impairmt factors :

- Number of imp.
- Length of imp.
- Interval between imp.

T1 : [Pitrey et al. SPIE OptEng 2010]

T3 : [Pitrey et al. IEEE EUVIP 2011]

Design of the common set

Constraints :

wide / balanced range of qualities

all kinds of distortion must be represented

Existing approaches :

- *conduct a dedicated experiment*
- *design common set with the experiments*

pros :

- data is well represented by the common set

cons :

- all exp. need to be known a priori
- low flexibility
- dedicated experiment

Our approach :

progressive construction of the common set

pros :

- no extra experiment needed
- flexible approach
- feasible in practice!

cons :

- part of each exp. dedicated to common set
- data is partly represented by common set

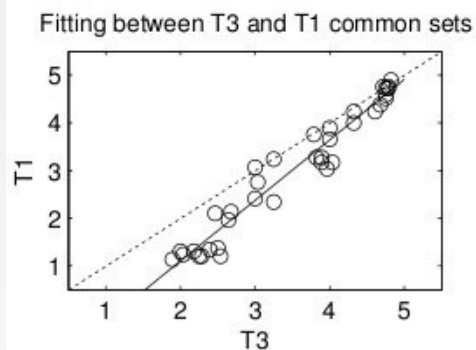
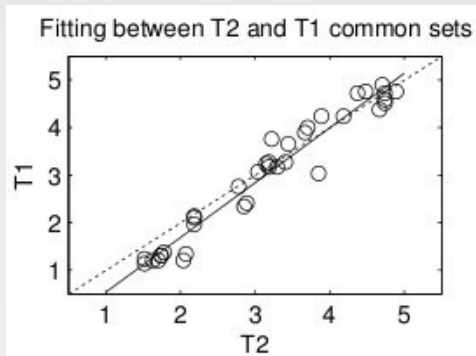
Choose a reference experiment

Common set between the 3 SVC experiments :

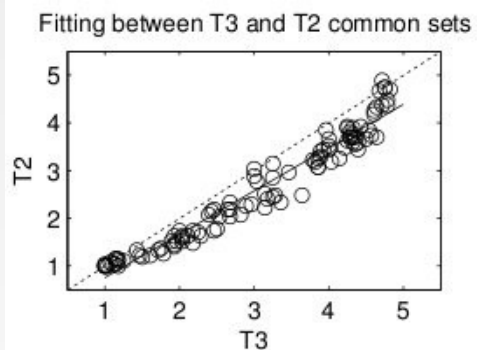
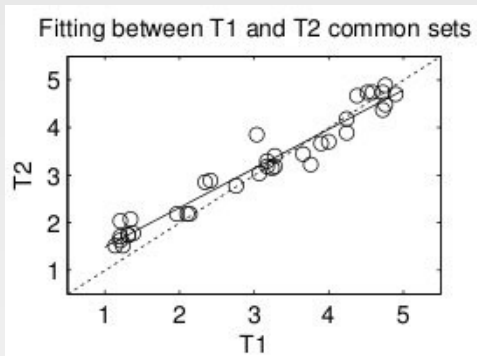
- Between T1 and T2 : 4 HRCs (36 sequences)
- Between T1 and T3 : 12 HRCs (108 sequences)
- Between T2 and T3 : 12 HRCs (108 sequences)

How to limit out-of-scale mapping?

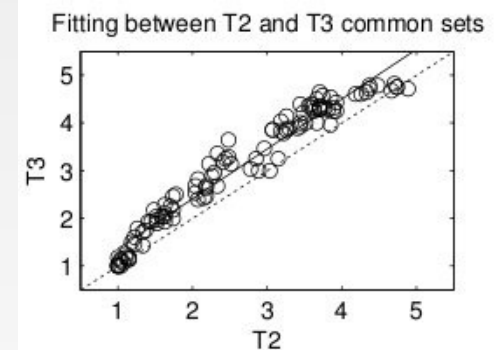
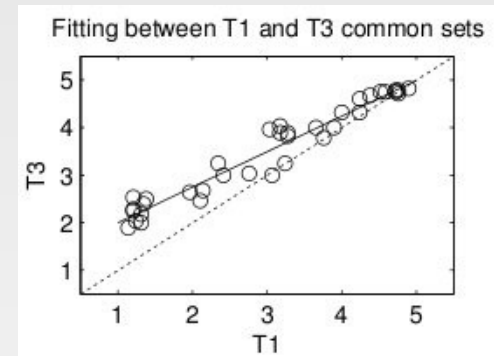
Mapping on T1



Mapping on T2



Mapping on T3

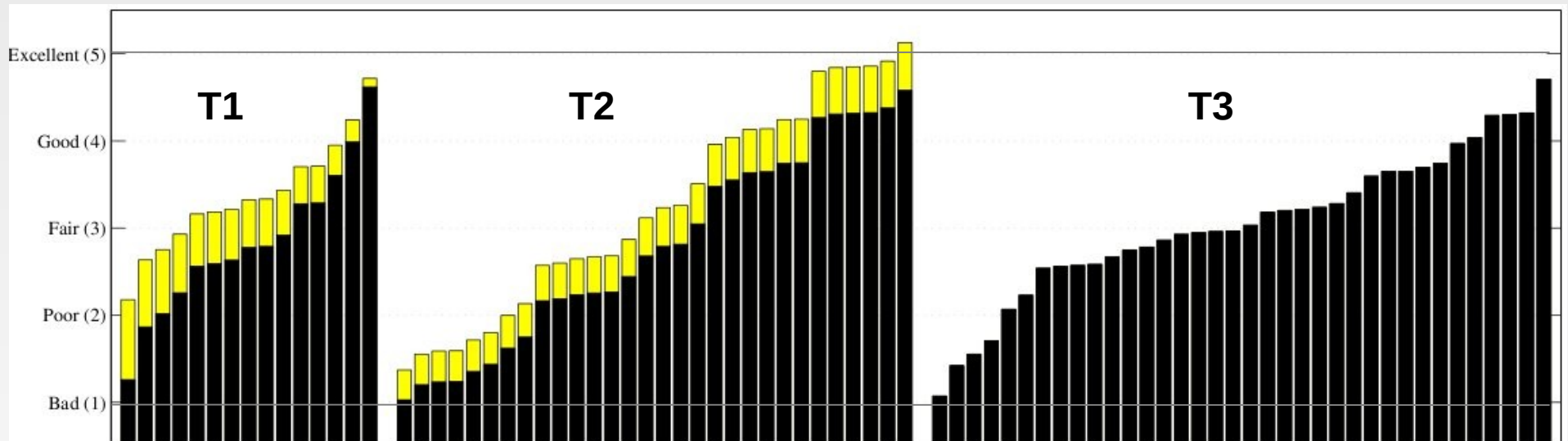


Mapping onto the reference experiment

Linear mapping onto the scale of T3 :

$$y_{T_i} = a_i x_{T_i} + b_i$$

MOS from each experiment before and after mapping :



Comparing experiments : 2 examples

T2 and T3 : upscaling VS transmission errors

one layer, 320x240, QP = 26 + upscale to 640x480

statistically equivalent to

2 SVC layers, QP = 44/32 + 1sec. impairment

T1 and T2 : constant bitrate VS constant QP

2 SVC layers, bitrate = 120/600 kbps + 1sec. impairment

statistically equivalent to

2 SVC layers, QP = 44/32 + 1sec. Impairment

Statistical equivalence with 95% probability using a student t-test.

Calculated on MOS after fitting.

Average over 9 video contents.

To take away

Inter-experiment comparison :

- Need for a common set between experiments
- Wide / balanced range of qualities

Our low cost approach :

- Construction of the common set along with the experiments
- No extra experiment needed
- Flexible and realistic approach
- Choose a reference test to limit out-of-scale fitted scores
- Linear fitting onto the scale of the reference experiment

Noticeable outcomes :

- Large set of subjective scores (T1+T2+T3 = 902 sequences)
- Different types of distortions compared

Open questions

- Different evaluation protocols (pair-comp. VS abs. cat. Rating)
- Limited evolution between experiments
- Mapping functions
- Determining the size of the common set automatically
- Validation of the fitting
 - Pearson's correlation might not be enough !
 - RMSE, Cross validation...

Thank you for your attention

Ingredients : detailed view

3 subjective experiments on SVC transmission

2009 → 2011
IRCCyN (SVC4QoE project)

Global configuration :
2 SVC layers : QVGA+VGA
Displayed video **always VGA@30fps**
Impairmt. simulation : **part of EL is lost**
BL upscale as error-concealment
9 video contents
ACR 5-level scale
Standard viewing conditions

TIMELINE

T1

SVC Error-concealment

- Constant bitrate :
- {120;200}, 600kbps
- 2 fps for BL :
- {15,30} fps
- 1-second network impairmt :
- $\frac{1}{4}$ of frame disappears
- 3 error-concealment methods :
- Always upscale BL
 - Switch entire frame
 - Patch part of frame

T2

SVC coding artifacts

- Constant QP : 16 HRCs
- {26;32;38;44} f. each layer
- 4 BL upscale HRCs :
- {26;32;38;44}

T3

imprmt distribution

- 2 constant QP HRCs :
- 44/32 , 38/32
- 3 impairmt factors :
- Number of imp.
 - Length of imp.
 - Interval between imp.

T1 : [Pitrey et al. SPIE OptEng 2010]
T3 : [Pitrey et al. IEEE EUVIP 2011]