SFB-TRR 161
Quantitative Methods for Visual Computing
Transregional Collaborative Research Center

Subjective Image Quality Assessment with Boosted Triplet Comparisons

Hui Men, Hanhe Lin, Mohsen Jenadeleh, Dietmar Saupe
University of Konstanz

May 10, 2022

Universität Stuttgart
Universität Konstanz
Max-Planck-Institut für biologische Kybernetik
DFG

Full Reference IQA Datasets

IQA Datasets	Distortion Levels
CID:IQ	5
CSIQ	3-5
LIVE	5-6
KADID-10k	5
TID2013	5
VCL@FER	6

Qualities range over the five ACR/DCR categories.

Are 5-6 distortion levels enough?

No: Content providers apply image compression, focus on the high quality range (approx. 1 JND).

To create datasets with 12 or more distortion levels up to 2 or 3 JND, we need to sharpen the subjective assessment (DCR).

Our proposal: **Boosted triplet comparisons**.

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Reference and Distorted Image

Ref.

Orig. Dist.

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Boosting (A)

$$v' = v_{ref} + \alpha(v_{dist} - v_{ref}) \quad (\alpha>1)$$

Ref.

Amplification (A)

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Boosting (A+Z)

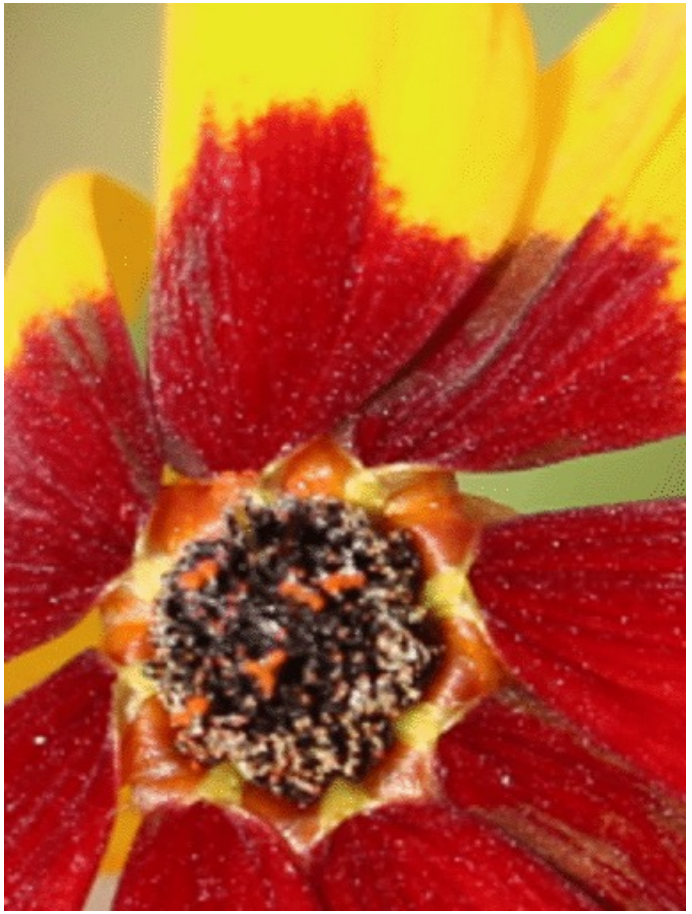
Ref.

Added Zoom (Z)

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Boosting (A+Z+F)

Ref.



Added Flicker (F)

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Triplet Comparisons (TC)

pivot

Which image is more similar to the middle one ?

left | not sure | right

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Thurstonian Scale Reconstruction from TC

$$Z_{ijk} = |X_k - X_j| - |X_i - X_j|$$

$$\begin{aligned}\Pr(Z_{ijk} > 0 \,|\, \boldsymbol{\mu}) &= 1 - \Phi(\mu_k - \mu_i) - \Phi\left(\frac{\mu_k + \mu_i - 2\mu_j}{\sqrt{3}}\right) \\ &\quad + 2\,\Phi(\mu_k - \mu_i)\,\Phi\left(\frac{\mu_k + \mu_i - 2\mu_j}{\sqrt{3}}\right) \\ \Pr(Z_{ijk} \le 0 \,|\, \boldsymbol{\mu}) &= 1 - \Pr(Z_{ijk} > 0 \,|\, \boldsymbol{\mu}).\end{aligned}$$

$$\begin{aligned}L(\boldsymbol{\mu}) &= -\sum_{(i,j,k,R_{ijk}) \in T} \log p^{R_{ijk}} (1-p)^{1-R_{ijk}} \\ p &= \Pr(Z_{ijk} > 0 \,|\, \boldsymbol{\mu}).\end{aligned}$$

$$\hat{\boldsymbol{\mu}} = \underset{\boldsymbol{\mu}=(\mu_0,\ldots,\mu_M)}{\arg\min} \; L(\boldsymbol{\mu}).$$

Algorithm 2 Probability of a response $R_{ijk} \in \{0,1\}$ to a triplet comparison (i,j,k)

$\boldsymbol{\mu} = (\mu_0, \ldots, \mu_M)$ ▷ stimuli means in model
$u_0 \leftarrow \mu_k - \mu_i$
$v_0 \leftarrow (\mu_k + \mu_i - 2\mu_j)/\sqrt{3}$
$p \leftarrow 1 - \Phi(u_0) - \Phi(v_0) + 2\Phi(u_0)\Phi(v_0)$
if $R_{ijk} = 1$ **then** ▷ stimulus i closer to j than k
    Return p
else ▷ stimulus k closer to j than i
    Return $1 - p$

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Three Experiments

1. Basic triplet comparisons
   (only the reference image as the pivot)

2. General triplet comparisons
   (arbitrary image as the pivot)

3. Boosting for degradation category rating (DCR)

Implemented by crowdsourcing on AMT.

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Materials

- 10 source images from the MCL-JCI Dataset

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Materials

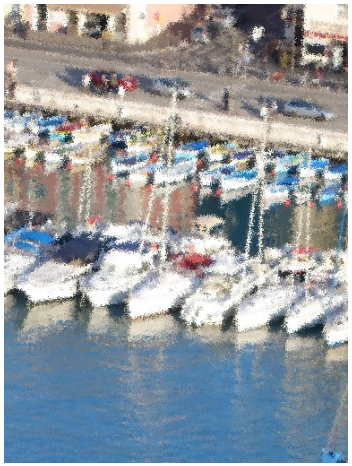
- 7 Distortion Types

Ref.

Color Diffusion

High Sharpen

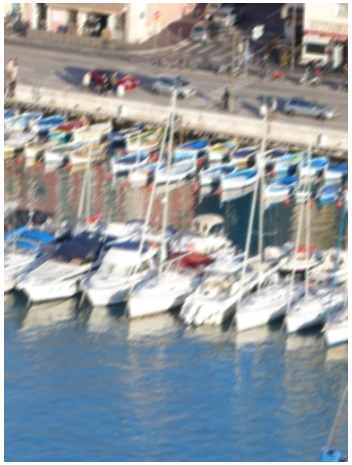
Jitter



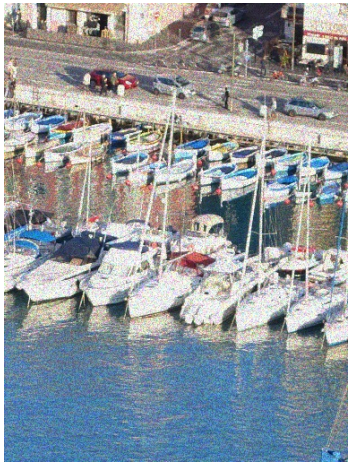
JPEG 2000

Lens Blur

Motion Blur



Multiplicative Noise



Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Experiment I (Baseline TC)

- 8 types of baseline TCs (pivot: reference image)
- 10 sources
- 7 distortion types
- **13** distortion levels (1 ref. + 12 dist.) ~ 3 JND
- Spacing between consecutive test images: **0.25** JND

Left Ref. Right

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Type 1/8 of Baseline TCs

- Plain TC (Original)

Left Ref. Right

Which image is more similar to the middle one ?

left not sure right

SFB-TRR 161

Type 2/8 of Baseline TCs

- A-Boosting (Amplified)

$$v' = v_{ref} + \alpha(v_{dist} - v_{ref}) \quad (\alpha=2)$$

Left Ref. Right

Which image is more similar to the middle one ?

left not sure right

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Type 3/8 of Baseline TCs

- Z-Boosting (cropped to 0.5x → Zoomed 2x)

Left Ref. Right

Which image is more similar to the middle one ?

left not sure right

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Type 5/8 of Baseline TCs

- AZ-Boosting (Amplified + Zoomed)

Left Ref. Right

Which image is more similar to the middle one ?

left not sure right

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Type 5/8 of Baseline TCs

- F-Boosting (Flicker)

Left (←→Ref.) Right (←→Ref.)

Which image has a stronger flicker effect?

left | not sure | right

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Type 6/8 of Baseline TCs

- AF-Boosting (Amplified + Flicker)

Left (←→Ref.) Right (←→Ref.)

Which image has a stronger flicker effect?

left not sure right

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Type 7/8 of Baseline TCs

- ZF-Boosting (Zoomed + Flicker)

Left (←→Ref.) Right (←→Ref.)

Which image has a stronger flicker effect?

left not sure right

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Type 8/8 of Baseline TCs

- AZF-Boosting (Amplified + Zoomed + Flicker)

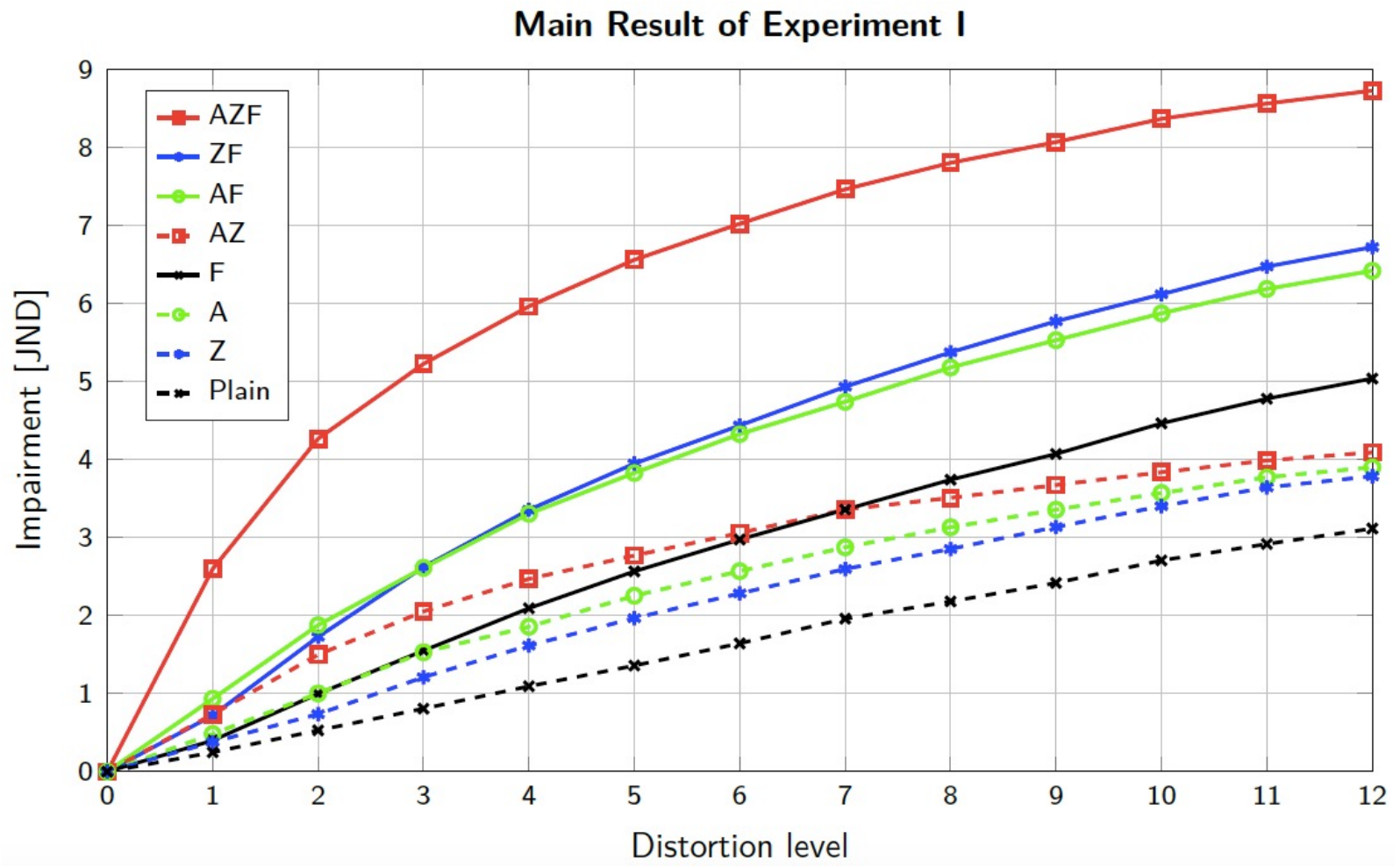
Left (←→Ref.) Right (←→Ref.)

Which image has a stronger flicker effect?

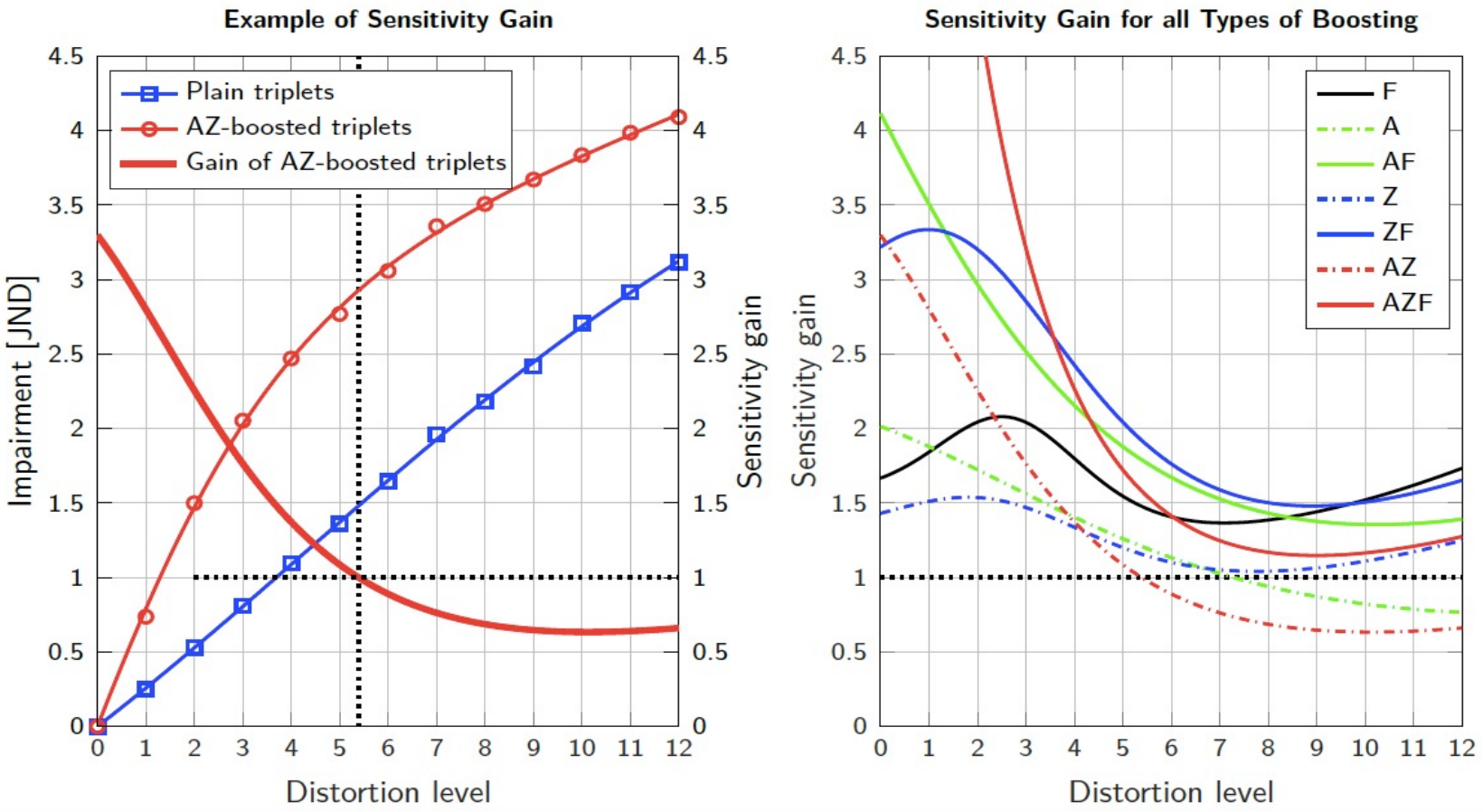
left not sure right

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

1. Reconstructed Impairment Scales



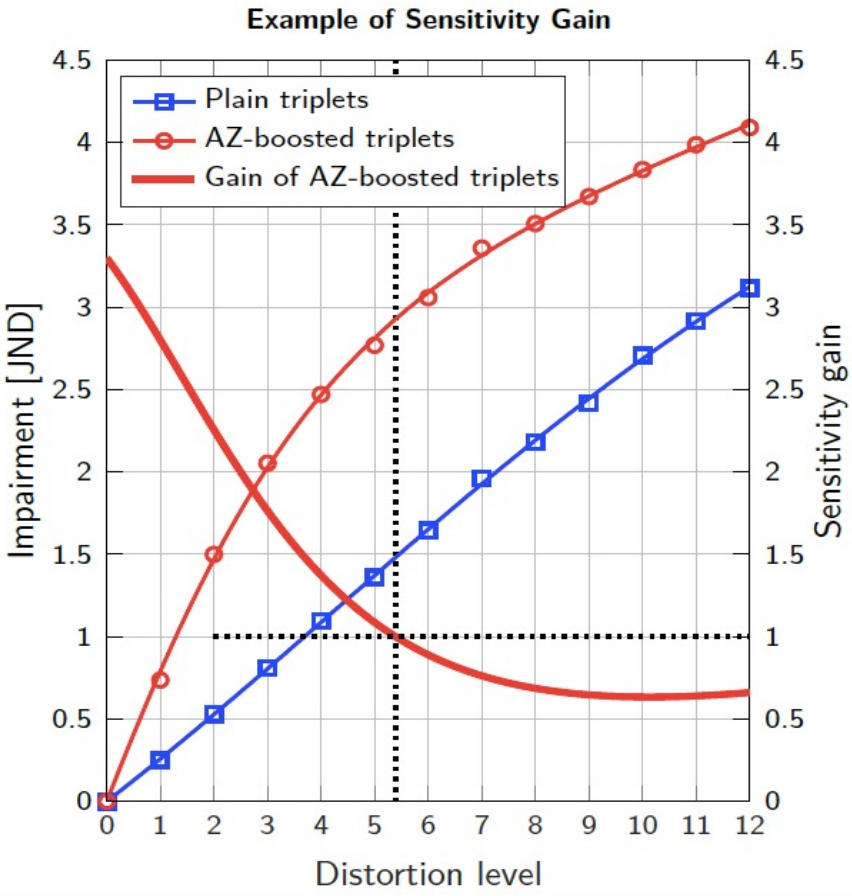
2. Sensitivity Gain

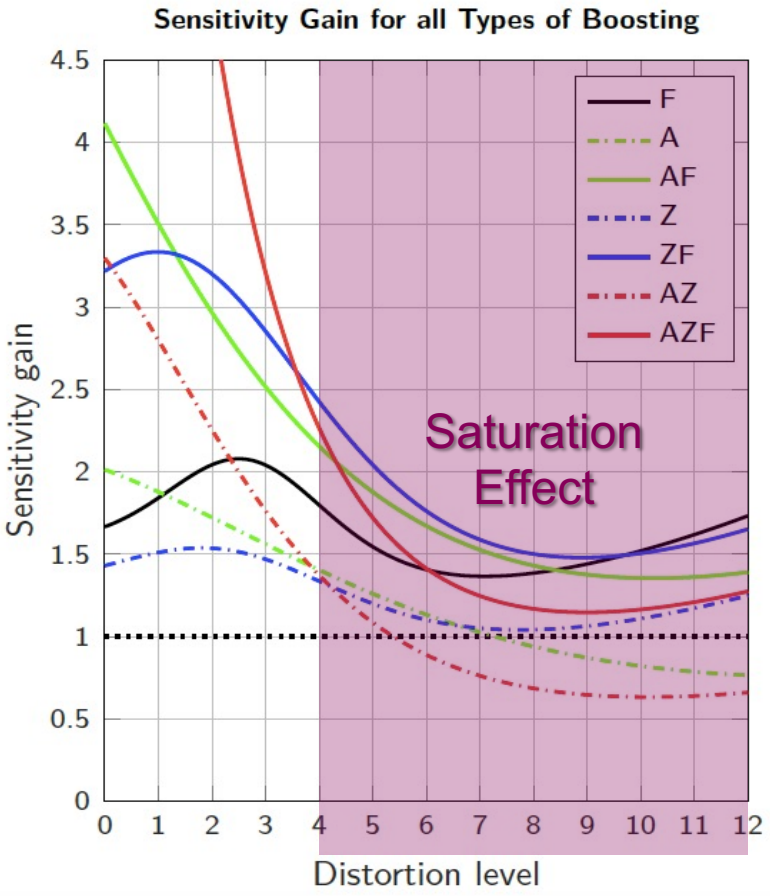


Sensitivity gain: Factor by which an increase of perceived distortion is multiplied by boosting.

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

2. Sensitivity Gain





Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Experiment II (General TC)

- 2 types of general TCs (pivot: distorted image)
  - plain
  - AZF-boosting
- 10 sources
- 1 distortion type (motion blur)
- **31** distortion levels (1ref. + 30 dist.) ~ 3 JND
- Spacing between consecutive test images: **0.1** JND

Left Dist. Right

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

2 Types of General TCs

- Plain TC (original)

Left Dist. Right

Which image is more similar to the middle one ?

left not sure not sure

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

2 Types of General TCs

- AZF-Boosting (Amplified + Zoomed + Flicker)

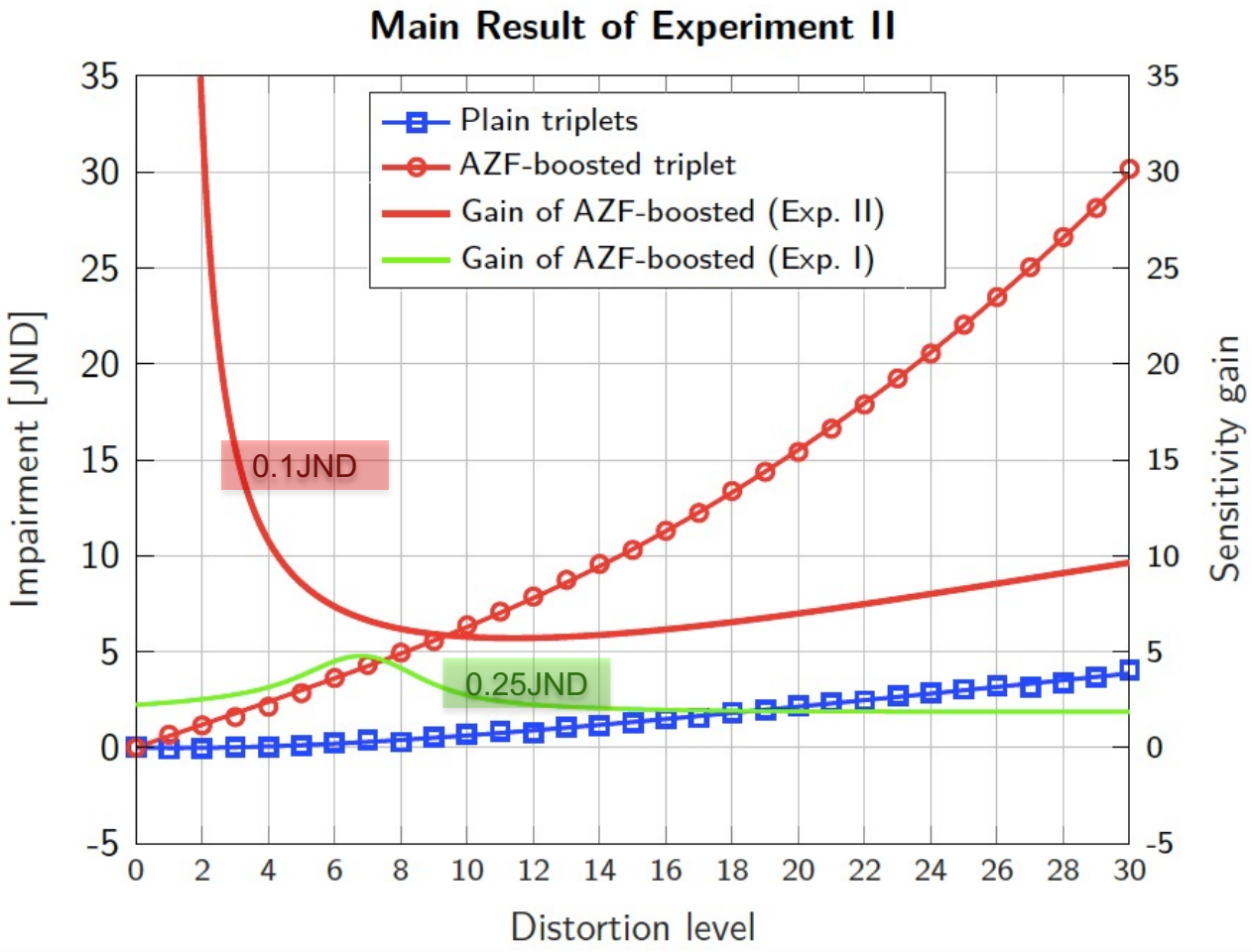
Left (←→Dist.) Right (←→Dist.)

Which image has a stronger flicker effect?

left not sure not sure

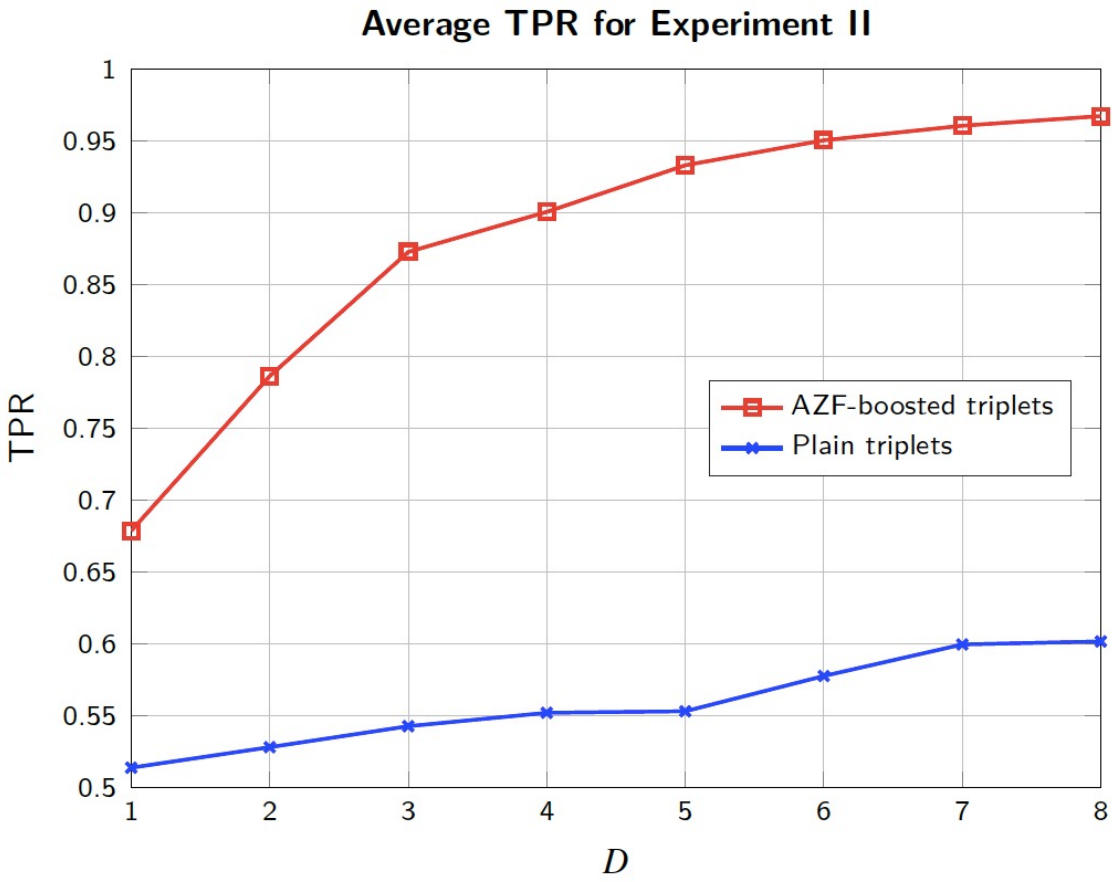
Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

1. Impairment Scales & Sensitivity Gain



Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

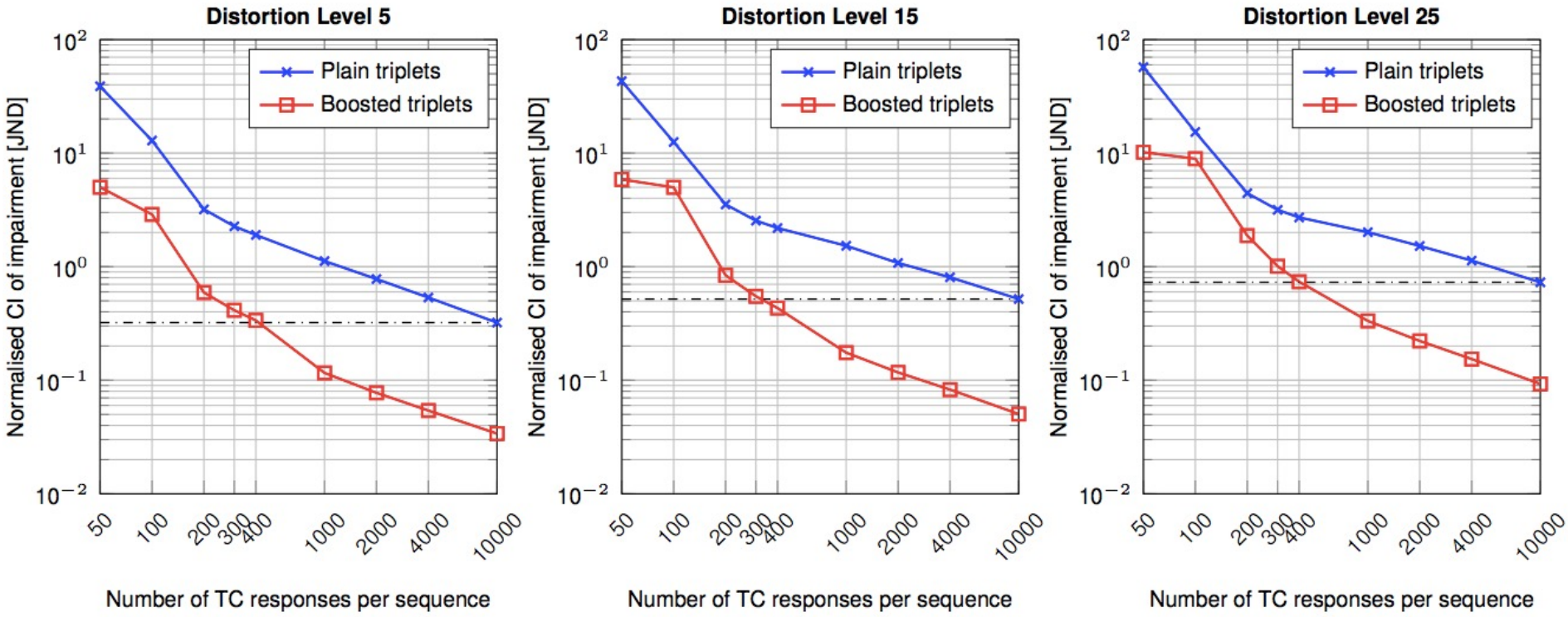
2. True Positive Rate



Average TPR for all triplets (i, j, k) and distance D = ||i-j|-|j-k|| for all 10 sources

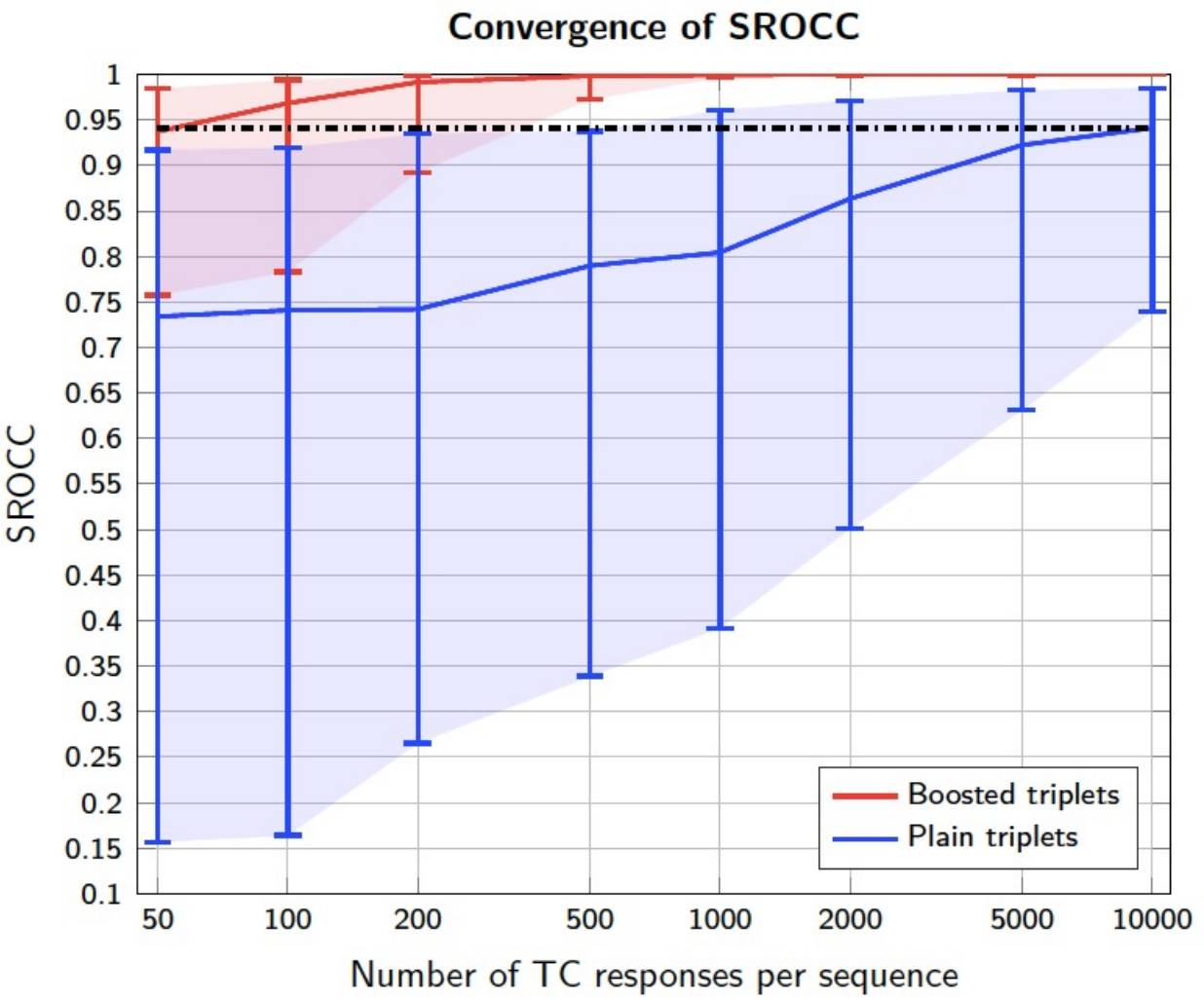
3. Convergence in Precision

- The precision of the reconstructions for given budgets of TCs
- 95% confidence intervals (CI)



Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

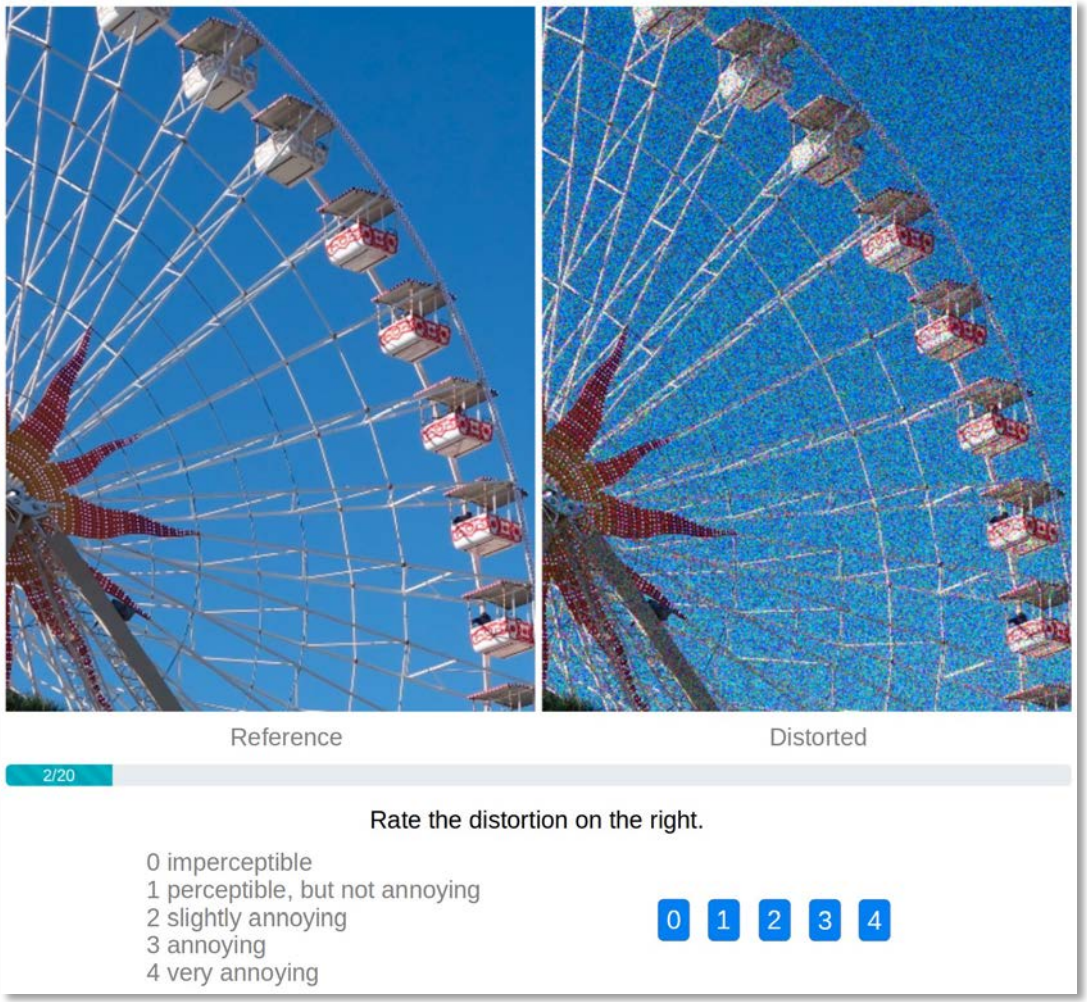
4. Convergence in Ordering



Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

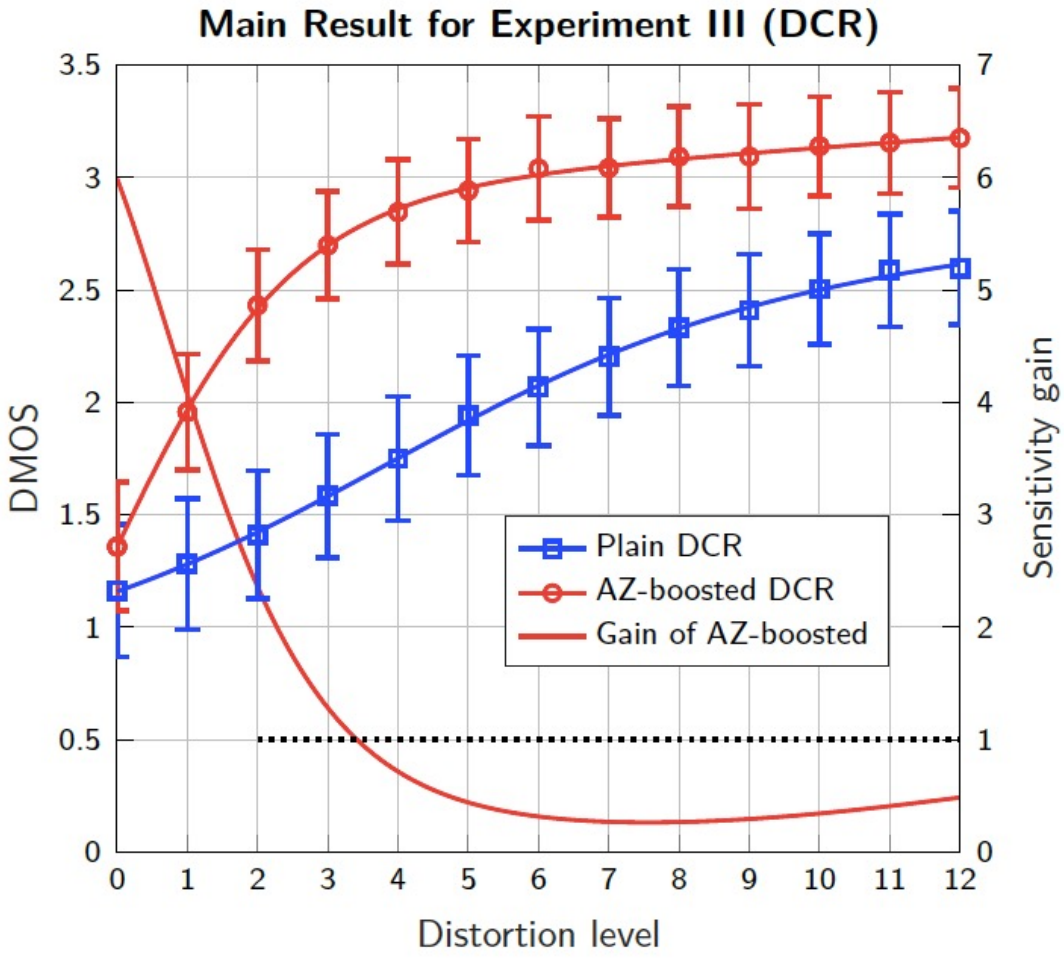
Experiment III (DCR)

- 4 types of DCRs (Plain, A, Z, AZ)
- 10 sources
- 7 distortion types
- 13 distortion levels
  (1 ref. + 12 dist.) ~ 3 JND
- Spacing between consecutive images: 0.25 JND



Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Result of Experiment III (DCR)

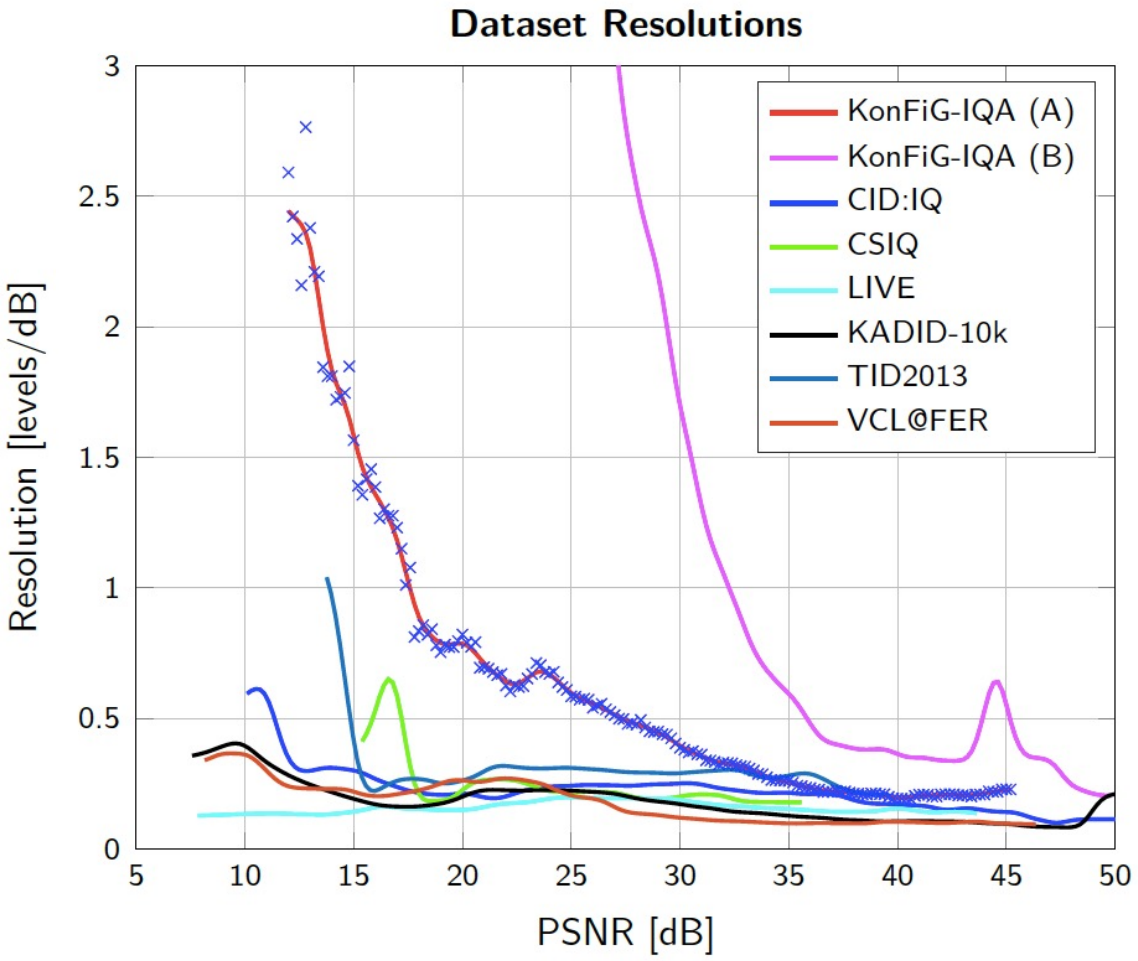


Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

KonFiG-IQA

- Konstanz **Fine-Grained** IQA Dataset

IQA Datasets	Distortion Levels
CID:IQ	5
CSIQ	3-5
LIVE	5-6
KaDiD-10k	5
TID2013	5
VCL@FER	6
KonFiG-IQA (A)	12
KonFiG-IQA (B)	30



Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021

Conclusion

1. Three boosting strategies: artefact amplification, zooming, flicker
   - enlarge the sensitivity of pair and triplet comparisons
   - increase the accuracy of subjective FR-IQA
2. Reconstruction of perceptual qualities from triplet comparisons
   - Thurstone‘s probabilistic model
   - maximum likelihood estimation
3. Two IQA datasets of 1140 images
   - 10 reference images, 7 dist. types, 12/30 dist. levels over 3 JND
   - 1.7 million crowdsourcing responses to triplet comparisons
4. Extensive FR-IQA performance analysis of boosted triplet comparisons
   - ratio of true positive responses
   - detection rates
   - sensitivity gains
   - effect sizes
   - convergence in accuracy
   - convergence in correlation
   - time complexity

Subjective image quality assessment with boosted triplet comparisons
IEEE Access, Vol. 9, pp. 138939-75, Oct. 2021