

Coherent rendering of smile previews with fast neural style transfer

Valentin Vasiliu^{1,2} Gábor Sörös^{2,3}



ISMAR'19, Beijing
October 15, 2019

video



AR smile previews

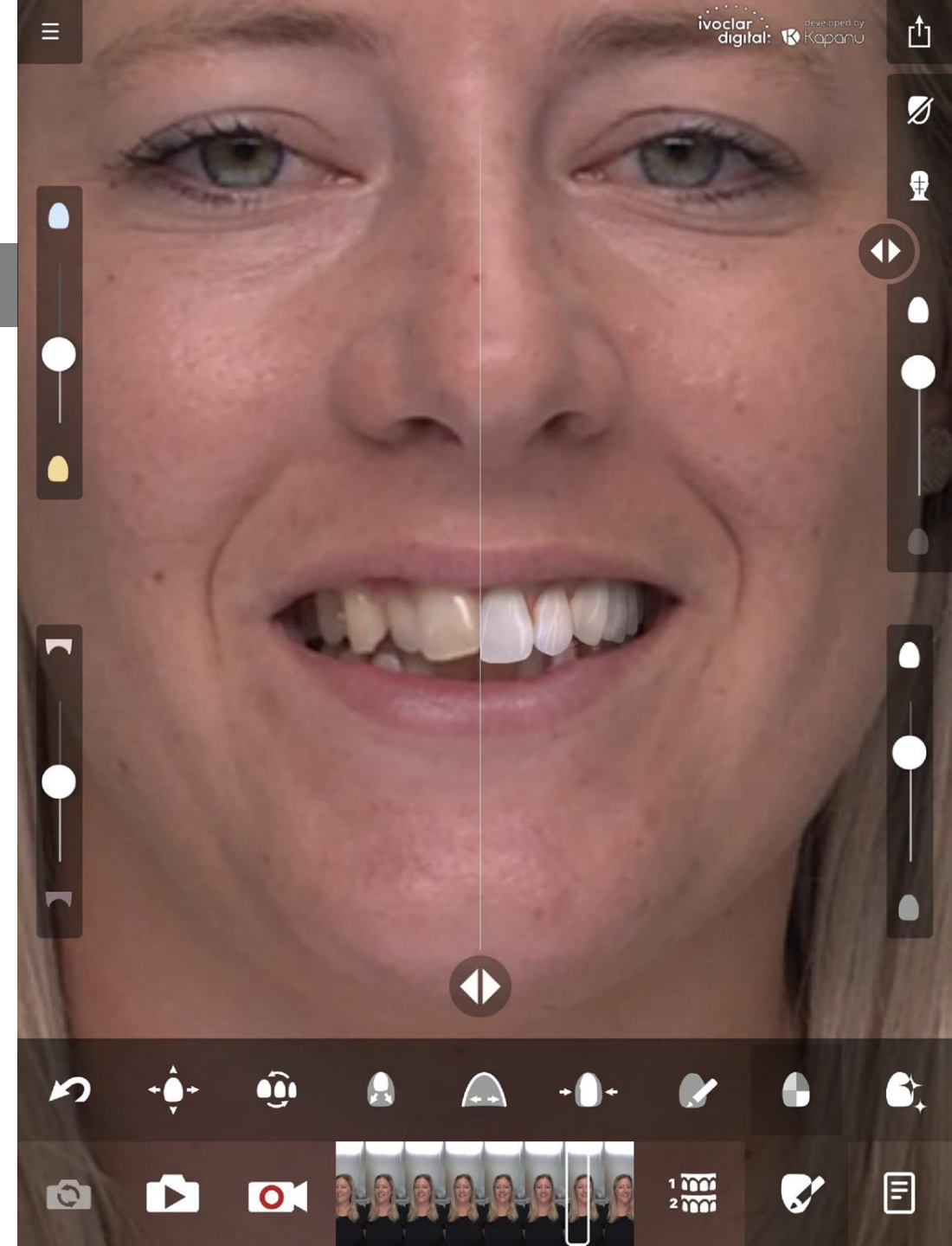
must be highly realistic

A „virtual mirror“:

- 3D and color
- live in motion
- mobile

Aesthetic visualization:

- increases awareness
- facilitates decision
- emotional



AR smile previews

must be highly realistic

Coherent rendering aims for seamless augmentation by modeling:

- color tone (illumination)
- defocus blur, motion blur
- camera lens distortions
- image noise, etc.

Capturing and modeling all these (dynamic) effects in the vision+graphics pipeline seems infeasible in real time.



negative examples

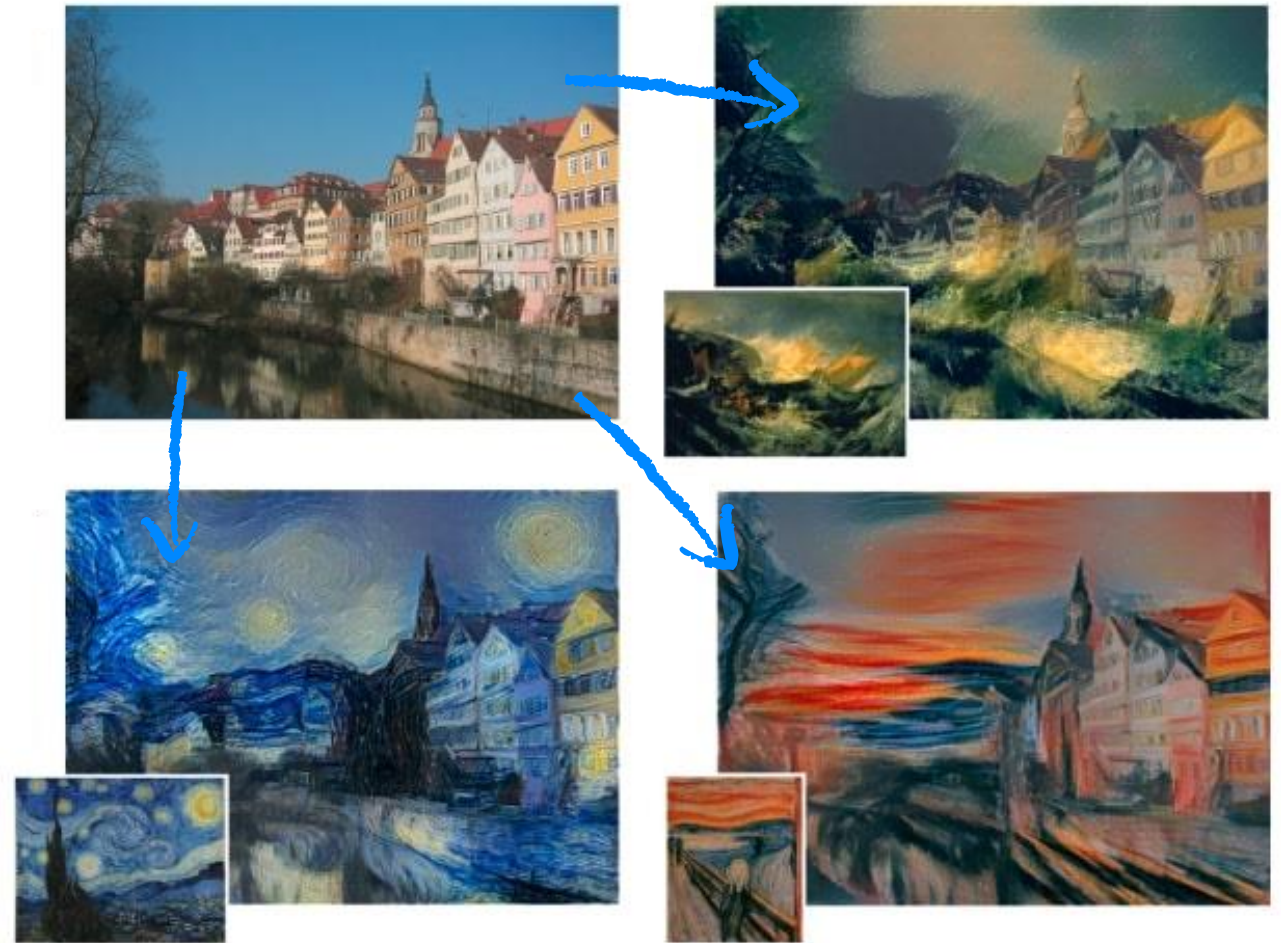


Neural style transfer

CNNs are able to separate the **content** of an image from the **style** of the image:

- filter banks
- hierarchical representation
- high-level features describe content
- low-level feature statistics describe style

By manipulating the statistics only, we can change the style of the image.



Style transfer as a post-processing step?



augmented image



original image

content image
→



generated by deepart.io
in a few minutes

→
style image

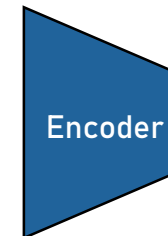
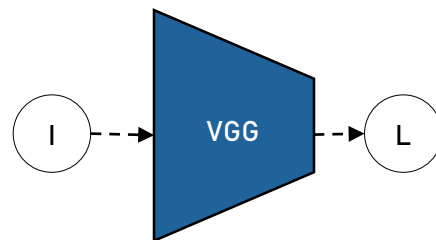
Additional challenges in AR:

- runtime
- no fix style
- temporal stability

Our goal: arbitrary style transfer with temporally stable results in < seconds time

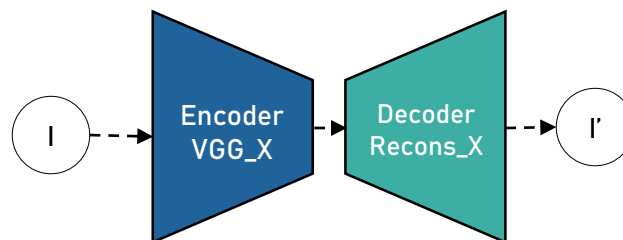
Network evolution

Image classifier



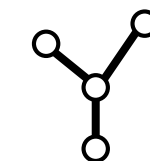
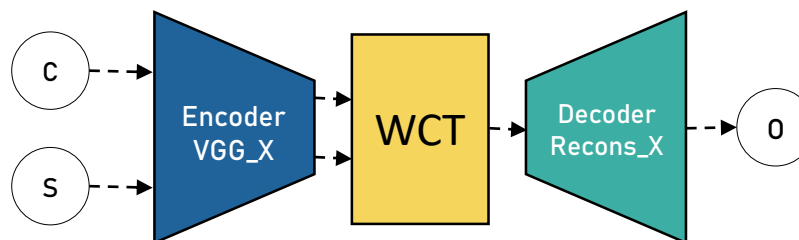
VGG19
MobileNet
MobileNetV2

Image autoencoder
X: truncation depth

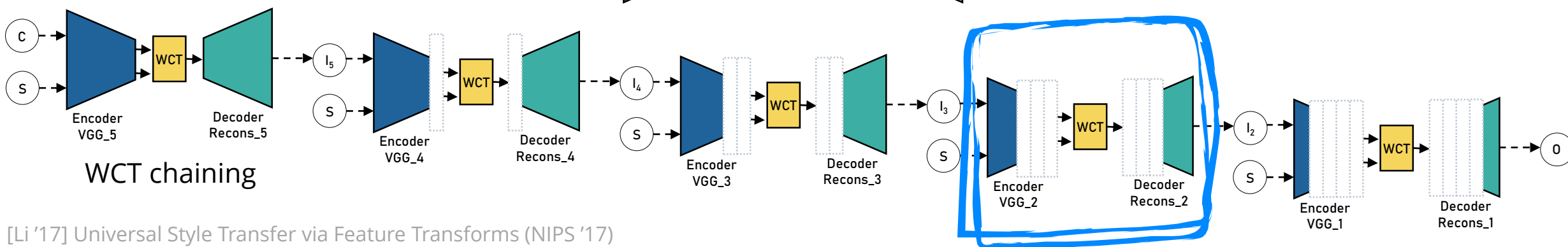


WCT [Li '17]
AdaIN [Huang '17]

WCT: Whitening-Coloring Transform
[Li'17] **only matrix operations**



truncation
chaining
U-Net

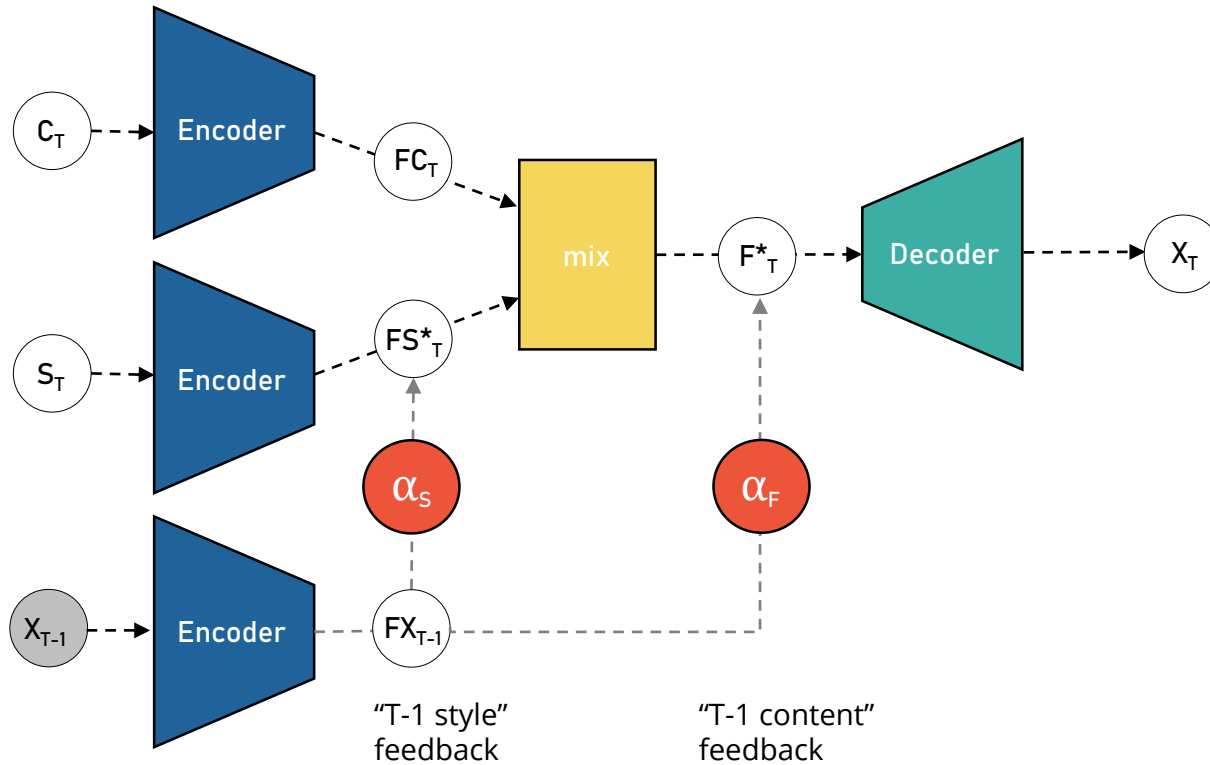


WCT chaining

[Li '17] Universal Style Transfer via Feature Transforms (NIPS '17)

[Huang '17] Arbitrary Style Transfer in Real-Time With Adaptive Instance Normalization (ICCV '17)

Network evolution (2)



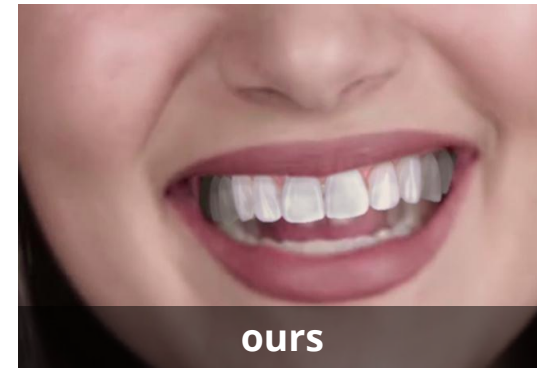
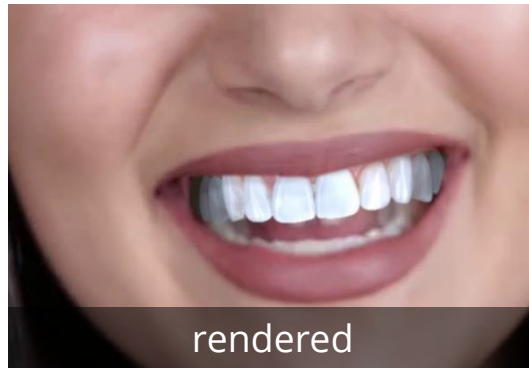
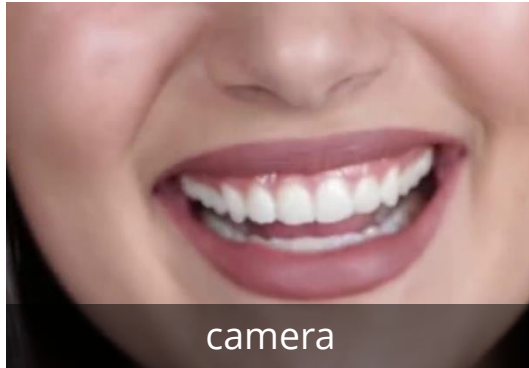
Training:

- only the decoder is trained
- reconstruction loss + style loss
- WCT: independent
- AdaIN: retraining with mixer once
- MS COCO images (no faces/teeth!)

Execution (PC):

- single forward pass
- VGG19.2 3s on CPU, 44ms on GPU

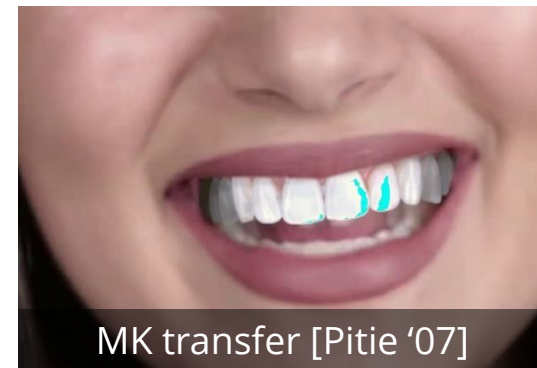
Comparison – color transfer



false colors



good



overflow

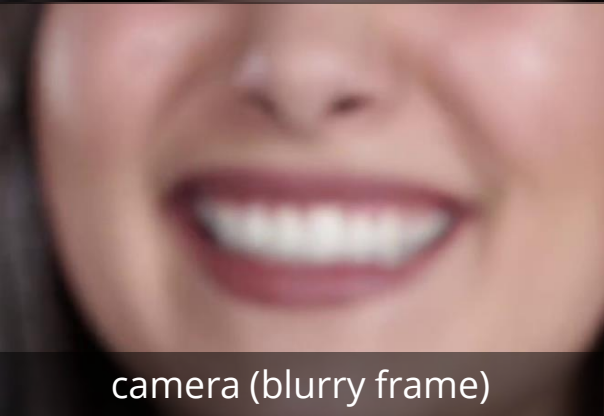
Quality improved: In our user study with 8 participants, everybody ranked all our results higher than the original rendering.

Comparison - blur

deeper architectures



rendered



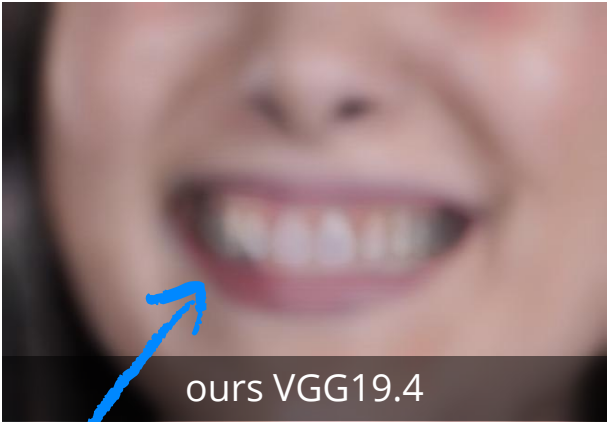
camera (blurry frame)



ours VGG19.1

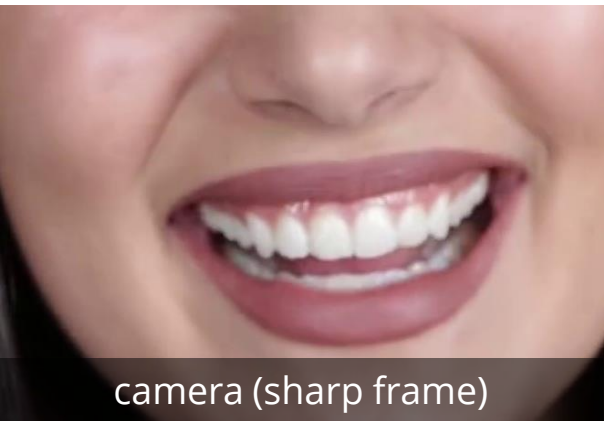


ours VGG19.2



ours VGG19.4

transfer blurriness well



camera (sharp frame)



ours VGG19.1

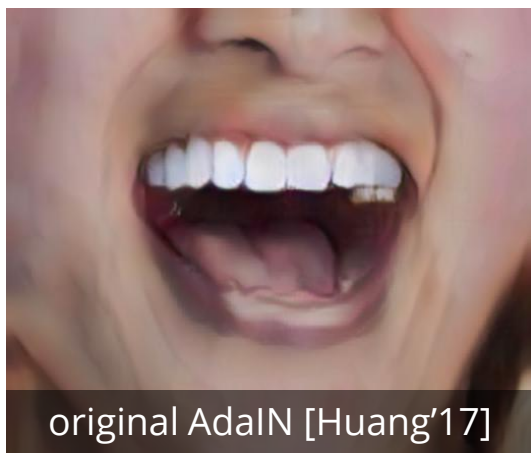
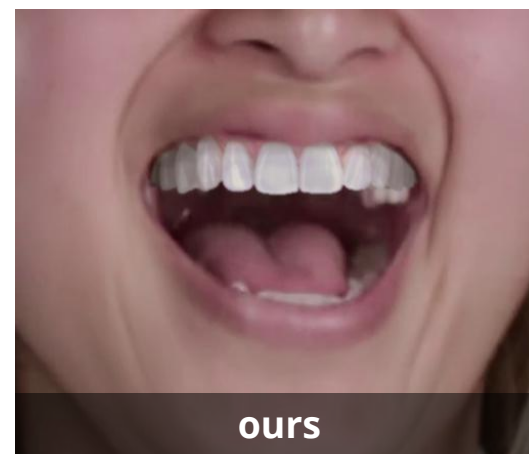
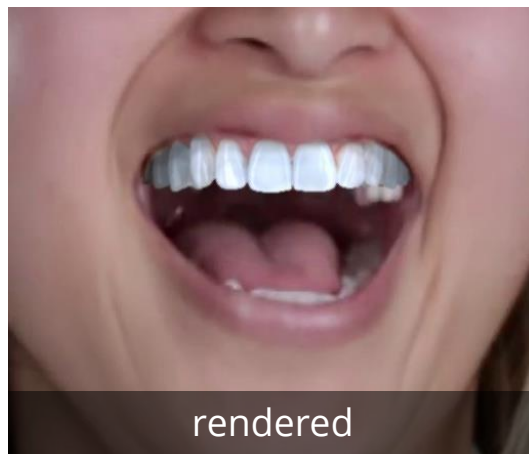
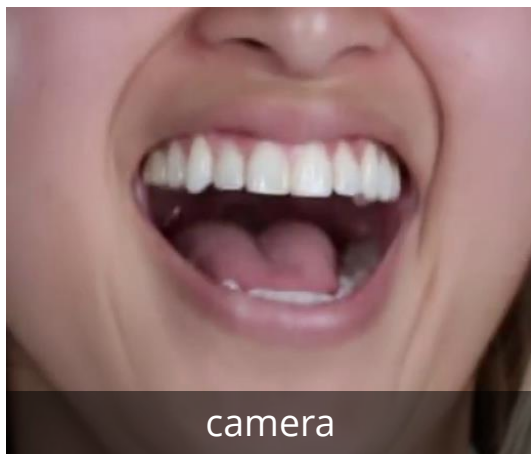


ours VGG19.2

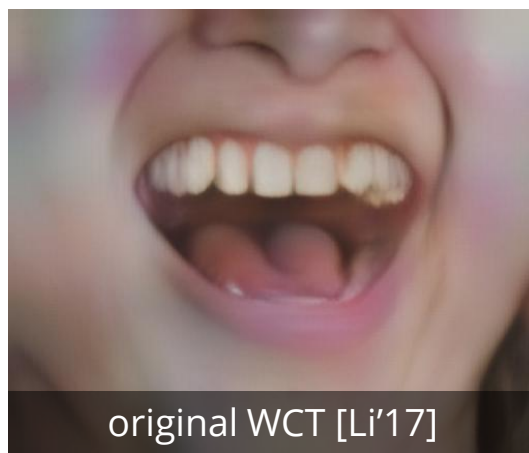


ours VGG19.4

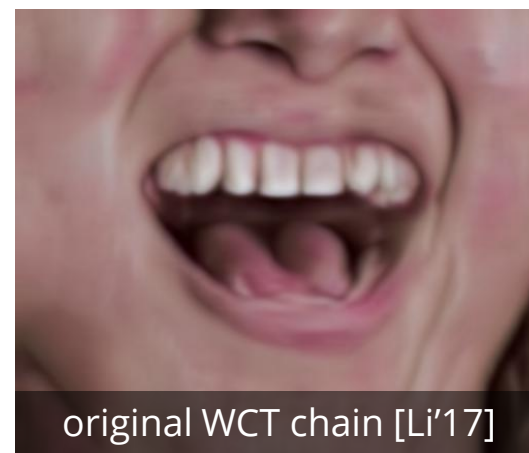
Comparison – original style transfer



too abstract



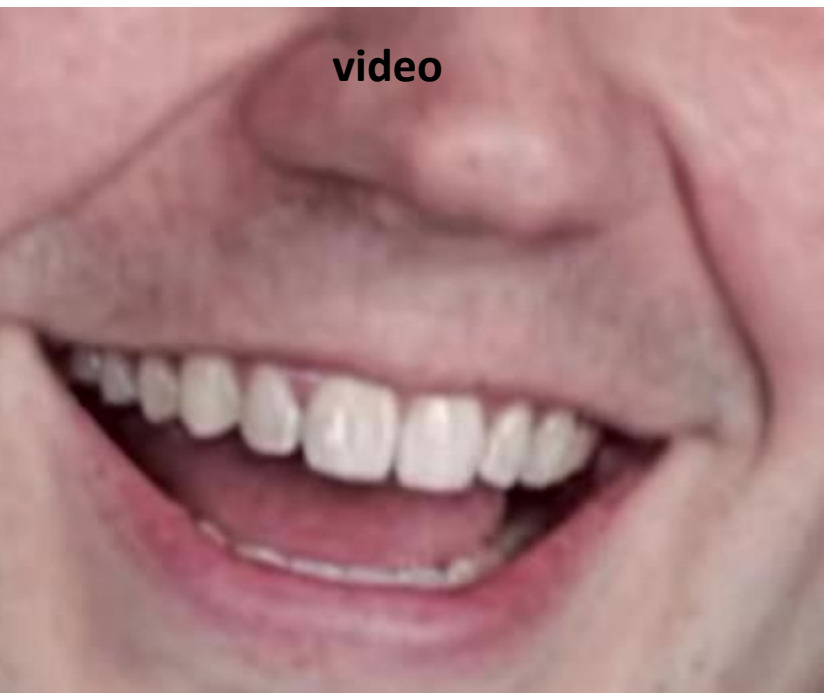
false colors, too abstract



too much content change



Comparison – temporal feedback



no feedback, colors somewhat flickering



correct feedback $\alpha_s=0.2, \alpha_F=0.1$



too much feedback $\alpha_s=0.5, \alpha_F=0.5$, ghosting

In case of shallower architectures, temporal feedback is not even necessary

Conclusion

Our method

improves the rendered content by transferring the style from the original background to the combined image (each new frame is a new style);

is only a post-processing step (no need to modify the AR pipeline);

re-generates the whole image in an autoencoder (instead of simple filtering);

can transfer color, image noise, and (to some extent) blur;

is not yet real time, but faster than other style transfer methods.

Further ideas

Targeted style transfer

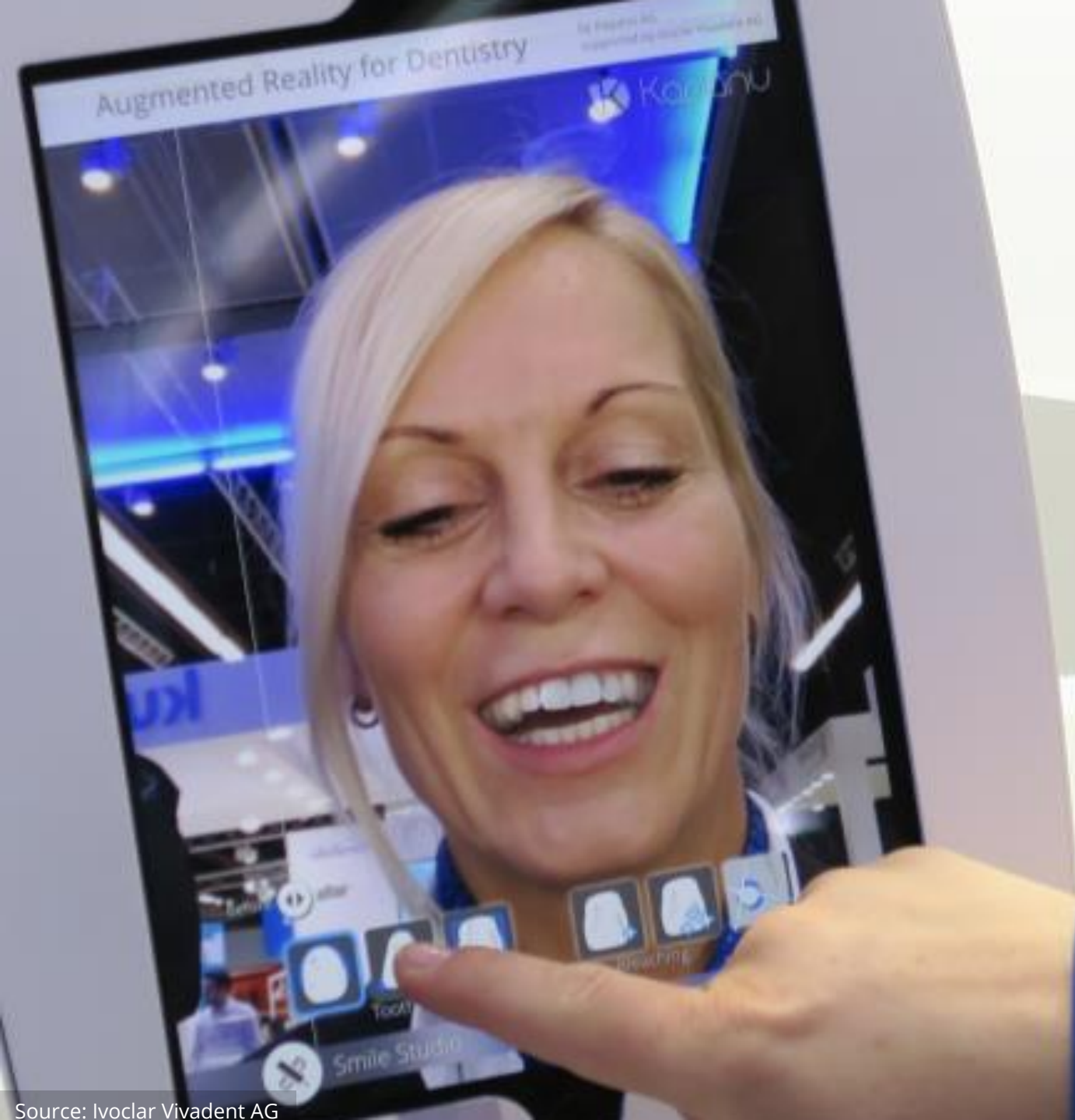
- binary masks for regional style transfer [Gatys '17]
- masks are cheap to get from the renderer
- further optimization of where to take the style patch from

Capturing other effects

- lens geometric distortions (part of the style?)

Other applications

- the method is not limited to faces or teeth, but we assumed we replace a real object with a virtual object of the same type and pose.



**Come
and try
our
demo!**

Thank You
www.kapanu.com

NOKIA Bell Labs



Sunnyvale

Murray Hill

Cambridge

Antwerp

Paris

Munich

Budapest

Madrid

Oulu

Espoo

Shanghai



internships around the globe
various AR/MR topics

<https://www.bell-labs.com/connect/internships/>

Thank You