ACCURATE COMPENSATION MAKES THE WORLD MORE CLEAR FOR THE VISUALLY IMPAIRED

Sijing Wu, Huiyu Duan, Xiongkuo Min, Danyang Tu, and Guangtao Zhai

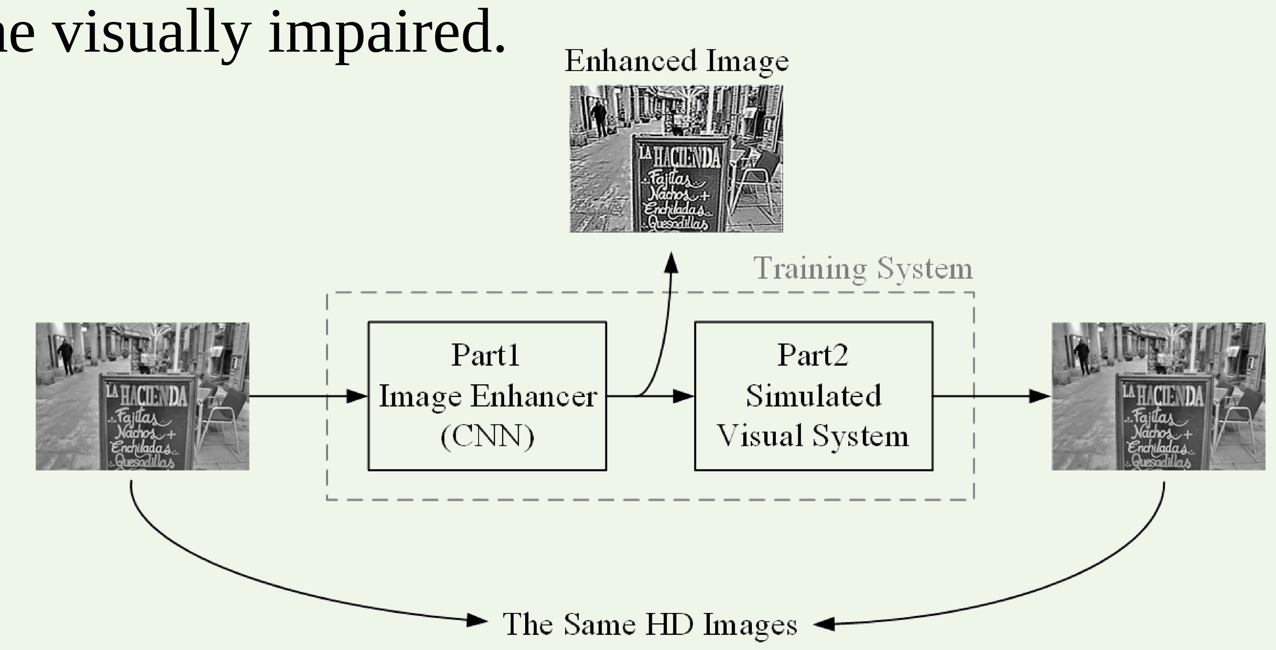
Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University, Shanghai, China

Background

- Visual impairment is one of the most serious social and public health problems in the world;
- Most of the visually impaired cannot be effectively assisted through conventional optical methods, while image enhancement is a good attempt;
- Image enhancement methods are also the foundation of assistive devices;
- Existing methods generally have some problems such as insufficient compensation and weak generality.

Image enhancement framework

Key idea: use a CNN to enhance the image, in order to compensate for the distortion caused by the visual system of the visually impaired.



The framework can be used to train image enhancement networks specialized for different visually impaired symptom that can be modeled.

References

[1] E. Peli et al, “Image enhancement for the visually impaired: simulations and experimental results,” in Investigative Ophthalmology & Visual Science 1991.

[2] Gang Luo et al, “Visual search performance of patients with vision impairment: effect of jpeg image enhancement,” in Ophthalmic and Physiological Optics 2012.

[3] Susana TL Chung and Gordon E Legge, “Comparing the shape of contrast sensitivity functions for normal and low vision,” in Investigative Ophthalmology & Visual Science 2016.

[4] Eli Peli and Tamar Peli, “Image enhancement for the visually impaired,” in Optical Engineering 1984.

Image Enhancement for Central Vision Loss

Using the proposed general deep learning based image enhancement framework for the visually impaired.

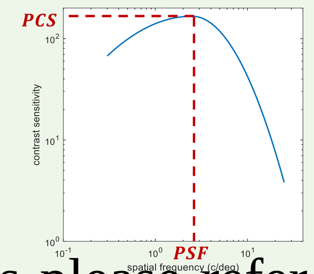
Visual system simulation

The loss of visual acuity and contrast sensitivity of the central vision loss can be simulated according to one’s contrast sensitivity function (CSF).

CSF approximation based on clinical indicators [3]:

$$CS_l(f_l)=\begin{cases} PCS_l-(f_l-PSF_l)^2 w_L^2 & if\ f<PSF \\ PCS_l-(f_l-PSF_l)^2 w_H^2 & if\ f\geq PSF \end{cases}$$

$$PCS=\frac{2-10^{-PR}}{10^{-PR}} \quad PSF_l=-10\log \mathrm{MAR}+\log_{10}(COF_N)-\frac{\sqrt{PCS_l}}{w_H}$$



where *PR* denotes the Pelli-Robson score. (details please refer to the paper.)

Perceptual image simulation pipeline [4]:

① Decompose the image into various spatial frequency bands;
② Calculate the local band-limited contrast in every band-pass filtered image;
③ Threshold them by the corresponding contrast detection threshold;
④ Merge the thresholded band-pass filtered images to get the simulated perceptual image.

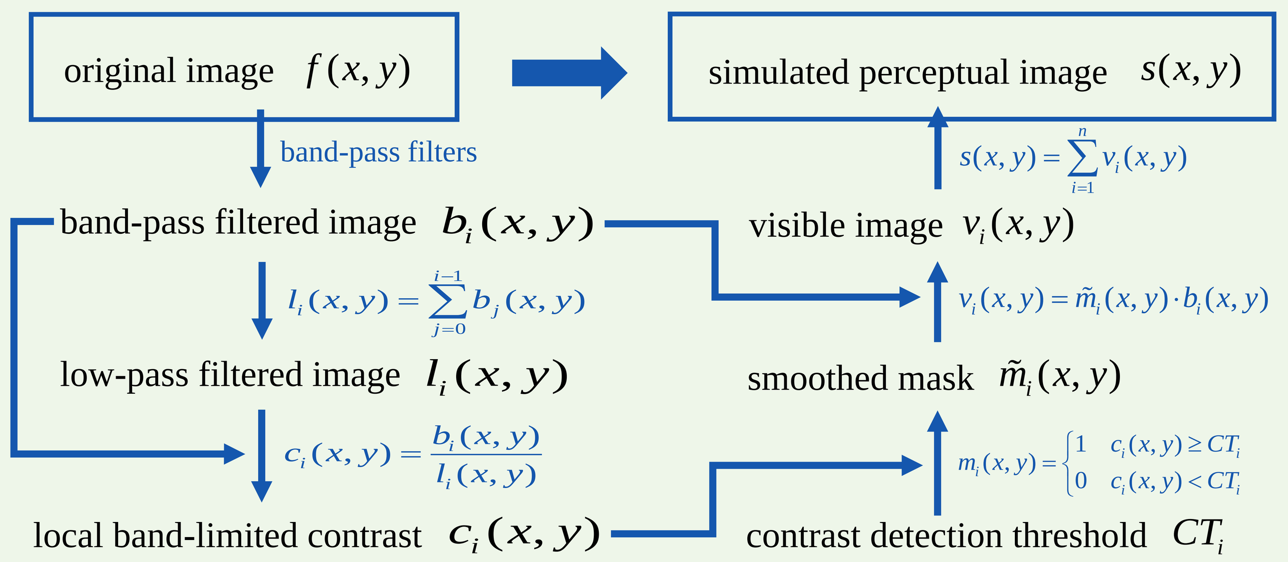


Image enhancement

- The image enhancement network is based on UNet structure (add BN between each convolutional layer and ReLU layer, and a sigmoid activation function in the last layer);
- Loss function: $L(\Theta)=\frac{1}{n}\sum_{i=1}^{n}\|\Psi(F(I_i;\Theta))-I_i\|^2$

  where $F(\Theta)$ denotes the CNN; Ψ denotes the visual system.

Acknowledgement

This work was supported in part by the National Natural Science Foundation of China under Grant 61901260, Grant 61831015, Grant 61771305, and Grant U1908210.

Experiment

Compared with two classic methods: adaptive enhancement [1] and enhancement in the DCT domain [2].

Simulation validation

By comparing the similarity between the simulated perceptual images of the enhanced images and the original images.

Quantitative result: Average PSNR, SSIM and MSE results on simulated perceptual images. (Best results are in bold.)

	Mild			Moderate			Severe		
	PSNR	SSIM	MSE	PSNR	SSIM	MSE	PSNR	SSIM	MSE
Original	24.494	0.7449	270.71	22.985	0.6683	376.37	18.720	0.5175	947.36
Adapt.[1]	18.945	0.7210	853.10	19.272	0.7154	794.48	18.901	0.5482	926.91
DCT.[2]	24.089	0.8077	283.40	22.889	0.7437	361.17	18.621	0.5177	969.25
Proposed	**24.530**	**0.8220**	**267.25**	**24.352**	**0.7950**	**308.98**	**20.852**	**0.5747**	**594.57**

Qualitative result: (a) is the original image. (b)(c)(d) are the enhanced images using adaptive enhancement [1], enhancement in the DCT domain [2] and the proposed method, respectively. (e)(f)(g)(h) are the simulated perceptual images of the severely visually impaired corresponding to (a)(b)(c)(d).

Patient experiment

15 myopic subjects with naked eyesight ranging from 0.1 to 0.8.

Objective part: aims to evaluate the improvement of the visual function of the visually impaired, which is tested by a searching task.

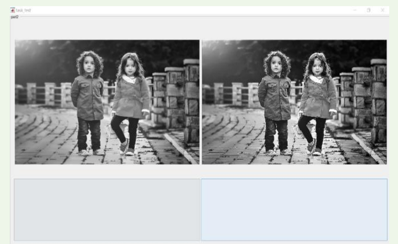
Mean and standard deviation of the accuracy in the searching task. Best result is in bold.

	Original	Adapt.[1]	DCT.[2]	Proposed
Accuracy (mean (±std))	0.822 (±0.137)	0.836 (±0.145)	0.853 (±0.124)	**0.893** (±0.094)

Subjective part: evaluates the improvement of the patient’s subjective perceptual quality by a comparison selection task.

Mean and standard deviation of the preference ratio.

	Original	Adapt.[1]	DCT.[2]
Preference Ratio (mean (±std))	0.757 (±0.202)	0.670 (±0.243)	0.757 (±0.184)



Preference ratio: the proportion that the subject prefers the images enhanced by the proposed method when comparing with another type of images.