Deep Automatic Portrait Matting

Xiaoyong Shen Xin Tao Hongyun Gao Chao Zhou Jiaya Jia

The Chinese University of Hong Kong

http://www.cse.cuhk.edu.hk/leojia/projects/automatting

Introduction

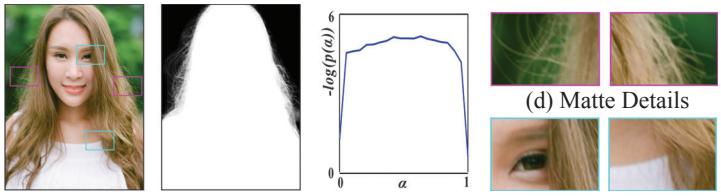
Problem

- Image Matting needs tedious user interactions
- The interactions are difficult to meet the algorithm requirement

Stokes Matting Trimap Matting

Challenges

- Learn automatic matting is very difficult
  - Rich matte details
  - Ambiguous semantic prediction
  - Discrepant matte value



(a) Input (b) Matte (c) Distribution (e) Semantic Regions

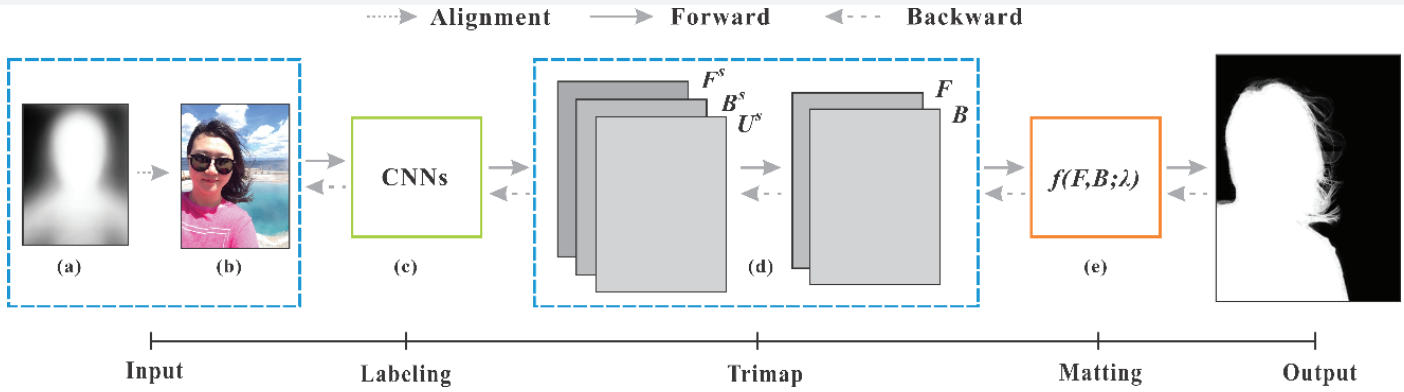
Contributions

- We proposed automatic portrait matting
  - An end-to-end deep matting CNNs framework
  - Novel matting layers
  - A matting dataset with 2,000 portraits

Dataset

- 2,000 portraits downloaded from Flickr
  - 1,700 for training and 300 for testing
  - Different age, gender, pose, hairstyle, background, camera type, etc.
  - The matting ground truth is estimated by human well labeled trimap

Our Framework



Trimap Labeling

- Input: RGB image
- Output: trimap representation
- Network: FCN [Long et al. 2015]

Image Matting Layer

- Input: trimap representation
- Output: alpha matte
- Newly-designed layers

- Feed-Forward

$$min\ \lambda A^T BA + \lambda(A-1)^T F(A-1) + A^T LA$$

- Back-Forward

$$\frac{\partial f}{\partial B} = -\lambda^2 D^{-1} diag(D^{-1}F), \frac{\partial f}{\partial F} = \frac{\partial f}{\partial B} + \lambda D^{-1}$$

$$\frac{\partial f}{\partial \lambda} = -\lambda D^{-1} diag(F+B) D^{-1} F + D^{-1} F$$

Experiments

Results of our deep automatic matting.