Active Matting

Xin Yang*, Ke Xu*, Shaozhe Chen, Shengfeng He, Baocai Yin, Rynson W.H. Lau
* joint first authors

NIPS
Neural Information Processing Systems

Introduction:

➢ What is Image Matting?

$$I_z = \alpha_z F_z + (1 - \alpha_z) B_z$$

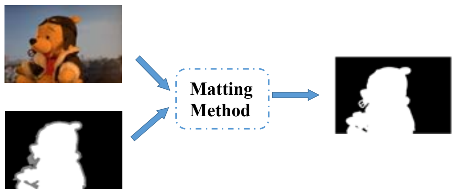
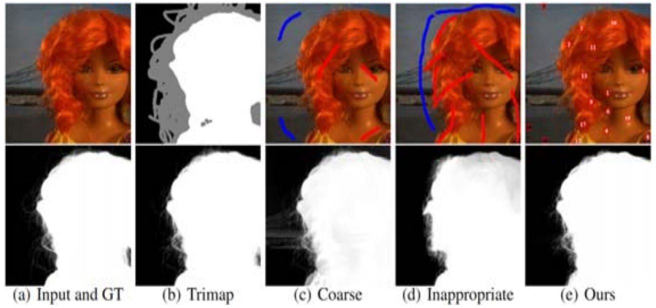


Image Matting (Alpha Matting) aims to solve the above equation, where $z = (x, y)$ denotes a pixel location in the input image I. F and B refer to the foreground and background images. α is the alpha matte whose values range between [0,1] defining the opacity of the foreground.

Matting Method takes a colorful image and its corresponding trimap as inputs and calculates an alpha matte.

➢ What are the limitations of conventional methods and what are the advantages of our proposed active model?



(a) Input and GT (b) Trimap (c) Coarse (d) Inappropriate (e) Ours

Conventional methods either require an accuracy trimap (Figure b), which is tedious to produce, or depend on the users' scribbles (Figures c and d), which rely on their expertise. In contrast, our active model is more user-friendly. It suggests proposals actively.

➢ What are our contributions?

- Propose an active model to detect informative regions for creating an alpha matte.
- Propose a recurrent network with a reinforcement learning strategy.
- Demonstrate that the learned informative knowledge can be generalized.

Methodology:

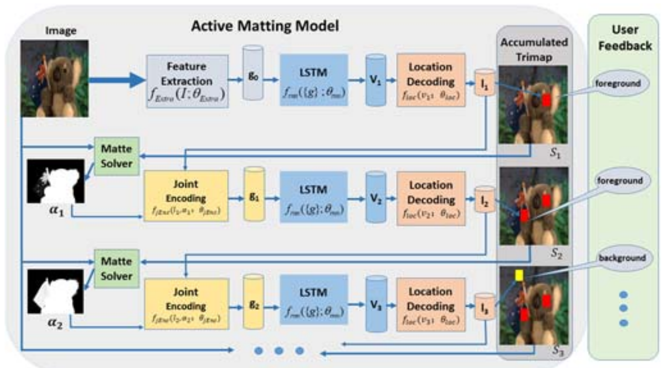
➢ Reinforcement learning of the sequence.

$$F = \sum_l p(l|I,\theta) \log p(\alpha|l,I,\theta) \qquad \frac{\partial F}{\partial \theta} = \sum_l p(l|I,\theta)\left[\log p(\alpha|l,I,\theta)\frac{\partial \log p(l|I,\theta)}{\partial \theta} + \boxed{\frac{\partial \log p(\alpha|l,I,\theta)}{\partial \theta}}\right]$$

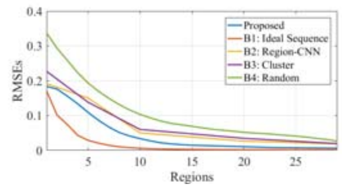
The accuracy is measured by Root-Mean-Square-Error (RMSE). Minimizing RMSE is equivalent to maximizing object function F. The training stage is by taking derivative of F *w.r.t* parameters θ. The second term in the *r.h.s.* of the derivative is dropped as the gradient is not available.

$$\frac{\partial F}{\partial \theta} = \frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{T}\left[\log p(\alpha|\hat{l}_i^m, I,\theta)\frac{\partial \log p(\hat{l}_i^m|I,\theta)}{\partial \theta}\right] \approx \frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{T}\left[(R_i^m - b_i)\frac{\partial \log p(\hat{l}_i^m|I,\theta)}{\partial \theta}\right] \quad R_i^j = \frac{\|\alpha_{i-1}-\alpha^{gt}\| - \left\|\alpha_i^j-\alpha^{gt}\right\|}{\max\{\|\alpha_{i-1}-\alpha^{gt}\| - \left\|\alpha_i^{\{S\}}-\alpha^{gt}\right\|\}}$$

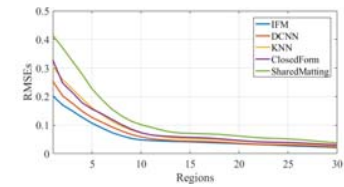
➢ Proposed Active Model



➢ Comparisons with baselines

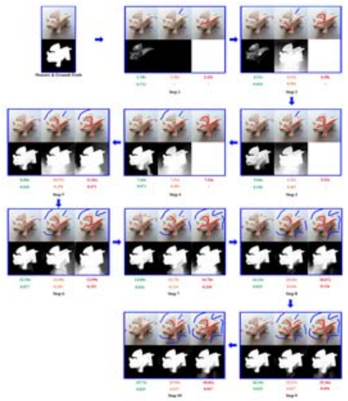


➢ Generalization results



Experiments:

➢ Qualitative results



➢ Quantitative results

	(image 1)	(image 2)	(image 3)	(image 4)	(image 5)	(image 6)	(image 7)	Mean RMSEs
Trimap+IFM	0.094 181s	0.046 209s	0.021 195s	0.018 192s	0.051 215s	0.017 187s	0.022 205s	0.031 204s
Ideal Sequence+IFM	0.096	0.048	0.023	0.019	0.064	0.019	0.022	0.032
Trimap+DCNN	0.117	0.049	0.021	0.019	0.079	0.019	0.021	0.033
Ideal Sequence+DCNN	0.121	0.051	0.022	0.022	0.081	0.021	0.021	0.034
Active+IFM	0.099 55s	0.051 66s	0.024 58s	0.019 62s	0.071 54s	0.021 55s	0.024 53s	0.034 57s
Active+DCNN	0.121	0.052	0.024	0.023	0.081	0.021	0.023	0.035
Trimap+Shared	0.104	0.059	0.023	0.021	0.071	0.024	0.022	0.035
Active+Shared	0.111	0.062	0.027	0.023	0.078	0.026	0.024	0.039
Trimap+KNN	0.097	0.053	0.029	0.031	0.066	0.022	0.037	0.043
Trimap+Global	0.111	0.059	0.024	0.023	0.072	0.018	0.024	0.045
Active+KNN	0.102	0.058	0.034	0.035	0.074	0.023	0.044	0.049
Trimap+Learning	0.126	0.107	0.028	0.015	0.084	0.022	0.015	0.051
Trimap+ClosedForm	0.136	0.117	0.029	0.016	0.097	0.023	0.016	0.056

References:

IFM: [Akosy et.al. CVPR' 2017] **Global**: [He et.al. CVPR' 2011] **KNN**: [Chen et.al. TPAMI' 2013]
Learning: [Zheng et.al. ICCV' 2009] **DCNN**: [Cho et.al. ECCV' 2016] **Shared**: [Gastal et.al. EG' 2010]