Anycost GAN

Рахманов Сергей, БПМИ192

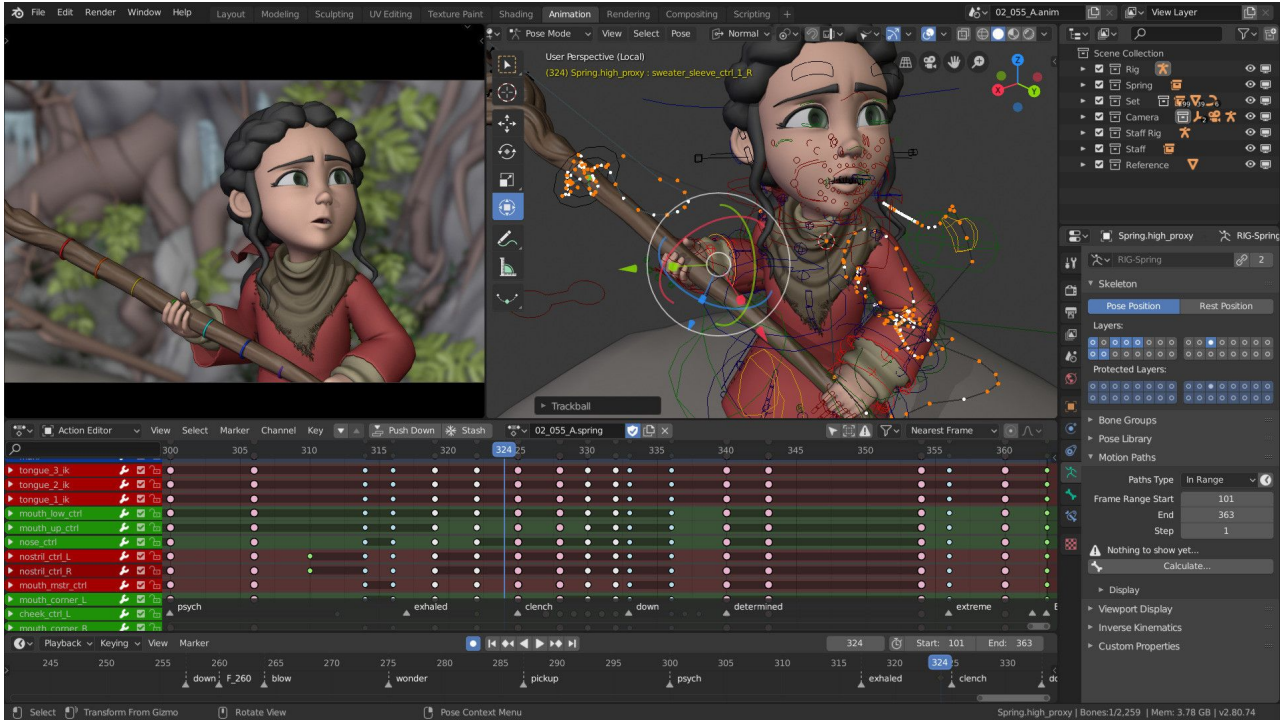
Вспомним StyleGAN

Вспомним StyleGAN

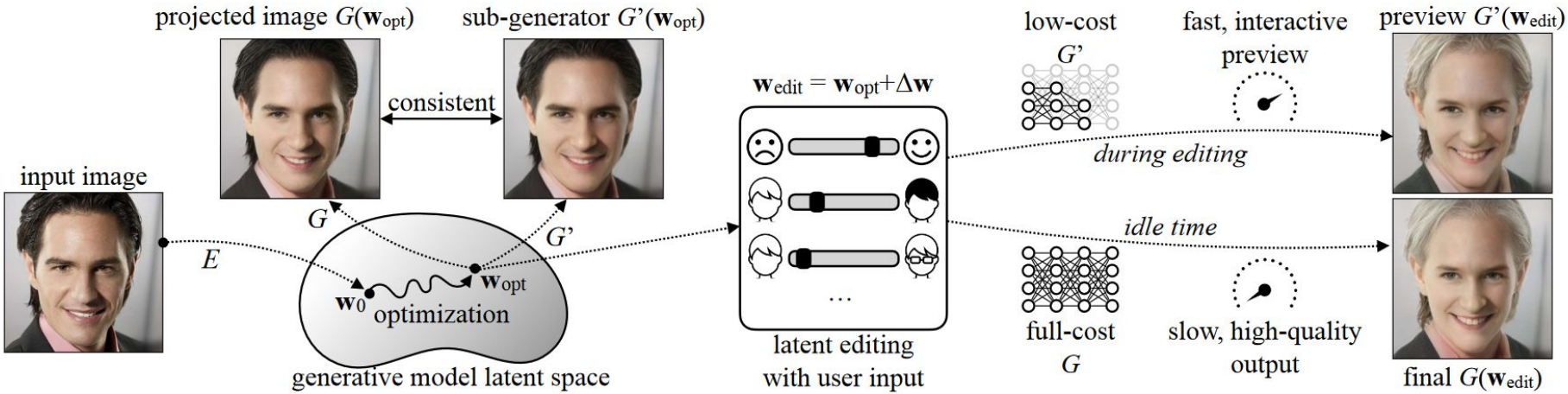
Вспомним StyleGAN

Красиво, но **дорого**!

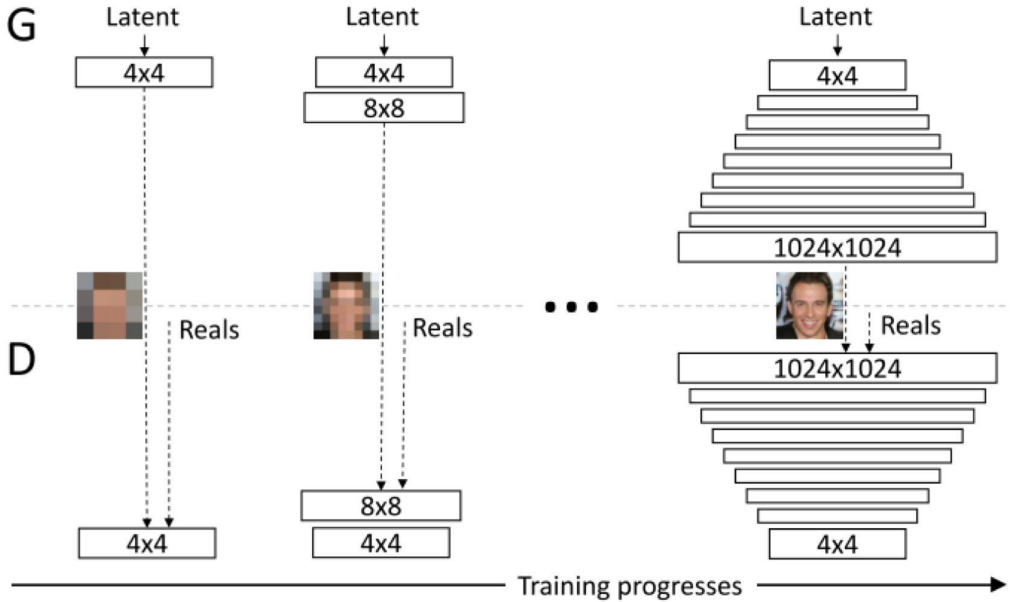
Идея


File Edit Render Window Help
Layout Modeling Sculpting UV Editing Texture Paint Shading Animation Rendering Compositing Scripting
Pose Mode View Select Pose
User Perspective (Local)
(324) Spring.high_proxy : sweater_sleeve_ctrl_1_R
Trackball
Scene Collection
Rig
Spring
Set
Camera
Staff Rig
Staff
Reference
Skeleton
Pose Position
Rest Position
Layers:
Protected Layers:
Bone Groups
Pose Library
Motion Paths
Paths Type
In Range
Frame Range Start
End
Step
Nothing to show yet...
Calculate...
Display
Viewport Display
Inverse Kinematics
Custom Properties
Action Editor View Select Marker Channel Key
Push Down
Stash
Nearest Frame
tongue_3_ik
tongue_2_ik
tongue_1_ik
mouth_low_ctrl
mouth_up_ctrl
nose_ctrl
nostril_ctrl_L
nostril_ctrl_R
mouth_mstr_ctrl
mouth_corner_L
cheek_ctrl_L
psych
exhaled
clench
down
determined
extreme
Playback Keying View Marker
Start: 101
End: 363
down
F_260
blow
wonder
pickup
psych
exhaled
clench
Select
Transform From Gizmo
Rotate View
Pose Context Menu

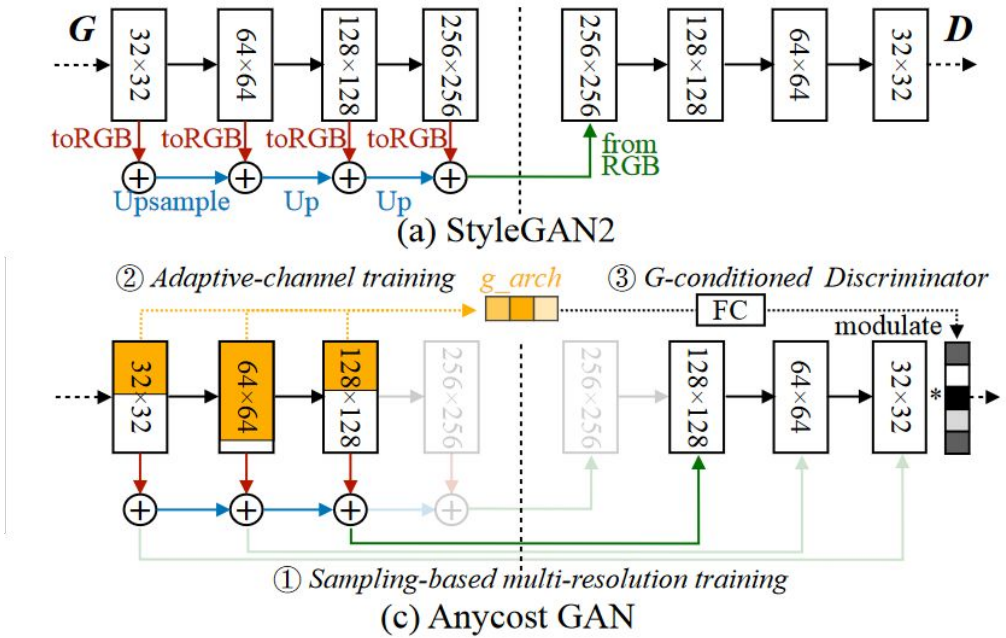
Идея


projected image $G(\mathbf{w}_{opt})$
sub-generator $G'(\mathbf{w}_{opt})$
consistent
input image
G
G'
E
$\mathbf{w}_0$
$\mathbf{w}_{opt}$
optimization
generative model latent space
$\mathbf{w}_{edit} = \mathbf{w}_{opt} + \Delta\mathbf{w}$
latent editing
with user input
low-cost
G'
fast, interactive
preview
during editing
idle time
full-cost
G
slow, high-quality
output
preview $G'(\mathbf{w}_{edit})$
final $G(\mathbf{w}_{edit})$

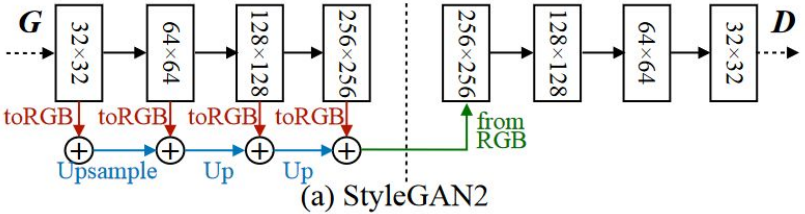
AnycostGAN: составной генератор



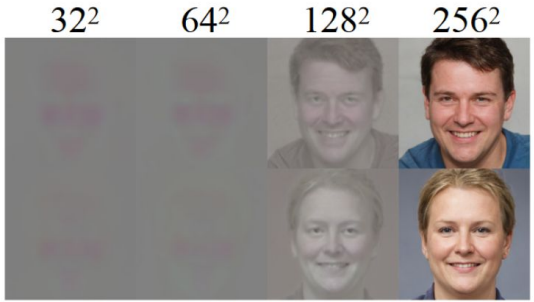
Составной генератор



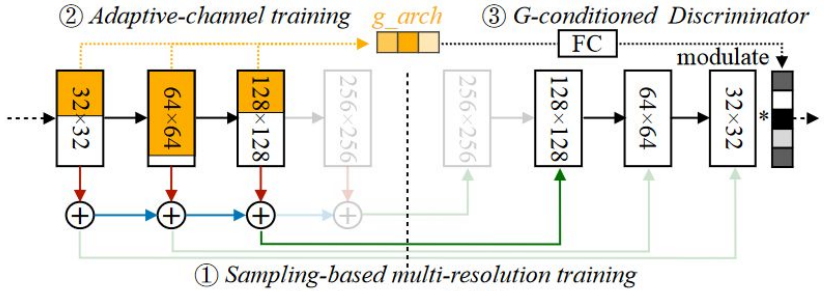
Составной генератор



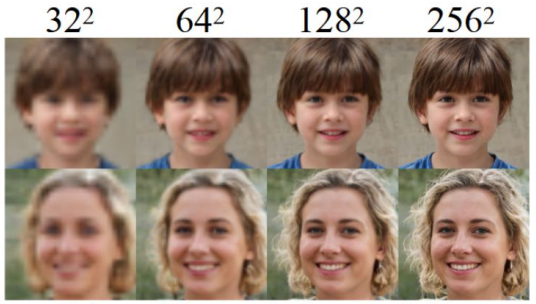
(a) StyleGAN2



(a) vanilla StyleGAN2

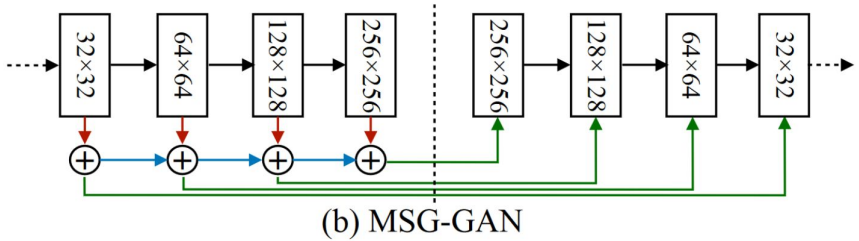


(c) Anycost GAN

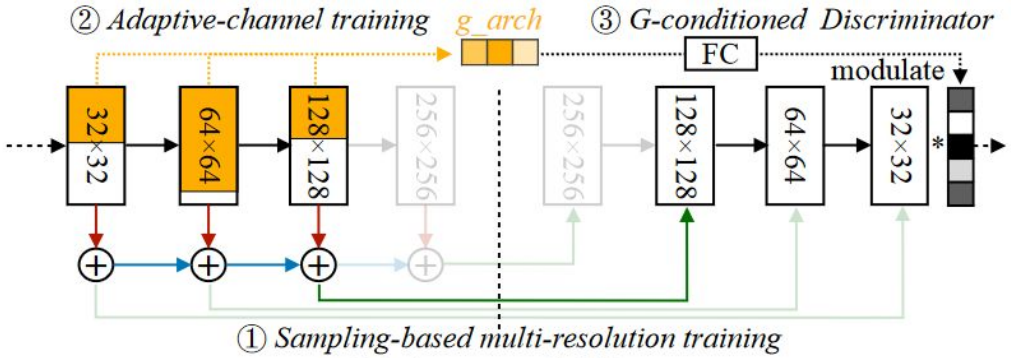


(b) multi-resolution training

Условный дискриминатор

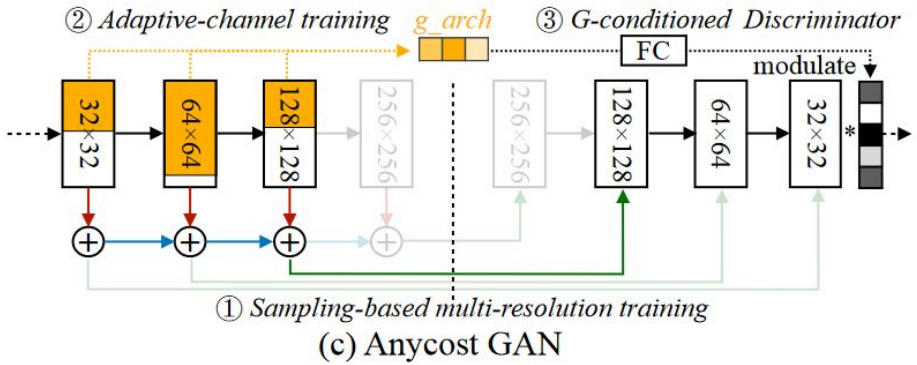


(b) MSG-GAN



(c) Anycost GAN

Адаптивные каналы



(c) Anycost GAN

$$\mathcal{L}_{\text{ada-ch}} = \mathbb{E}_{\mathbf{x},k}[\log D(\mathbf{x}^k)] + \mathbb{E}_{\mathbf{w},k,\mathbb{C}}[\log(1 - D(G^k_{\mathbb{C}}(\mathbf{w})))]$$

Сохраняем консистентность

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{ada-ch}} + \mathbb{E}_{\mathbf{w},k,\mathbb{C}}[\ell(G^k_{\mathbb{C}}(\mathbf{w}), G(\mathbf{w}))],$$

(c) w/o consistency loss (d) w/ consistency loss

Использование

Проецирование:

$$E^* = \arg\min_E \mathbb{E}_{\mathbf{x}}[\ell(G(E(\mathbf{x})), \mathbf{x}) + \alpha \mathbb{E}_{k,\mathbb{C}} \ell(G^k_{\mathbb{C}}(E(\mathbf{x})), \mathbf{x})]$$

$$\mathbf{w}^* = \arg\min_{\mathbf{w}}[\ell(G(\mathbf{w}), \mathbf{x}) + \alpha \mathbb{E}_{k,\mathbb{C}} \ell(G^k_{\mathbb{C}}(\mathbf{w}, \mathbf{x})]$$

Использование

Проецирование:

$$E^* = \arg\min_E \mathbb{E}_{\mathbf{x}}[\ell(G(E(\mathbf{x})), \mathbf{x}) + \alpha \mathbb{E}_{k,\mathbb{C}} \ell(G^k_{\mathbb{C}}(E(\mathbf{x})), \mathbf{x})]$$

$$\mathbf{w}^* = \arg\min_{\mathbf{w}}[\ell(G(\mathbf{w}), \mathbf{x}) + \alpha \mathbb{E}_{k,\mathbb{C}} \ell(G^k_{\mathbb{C}}(\mathbf{w}, \mathbf{x})]$$

Редактирование:

$$G(\mathbf{w} + \Delta\mathbf{w})$$

$$G^k_{\mathbb{C}}(\mathbf{w} + \Delta\mathbf{w})$$

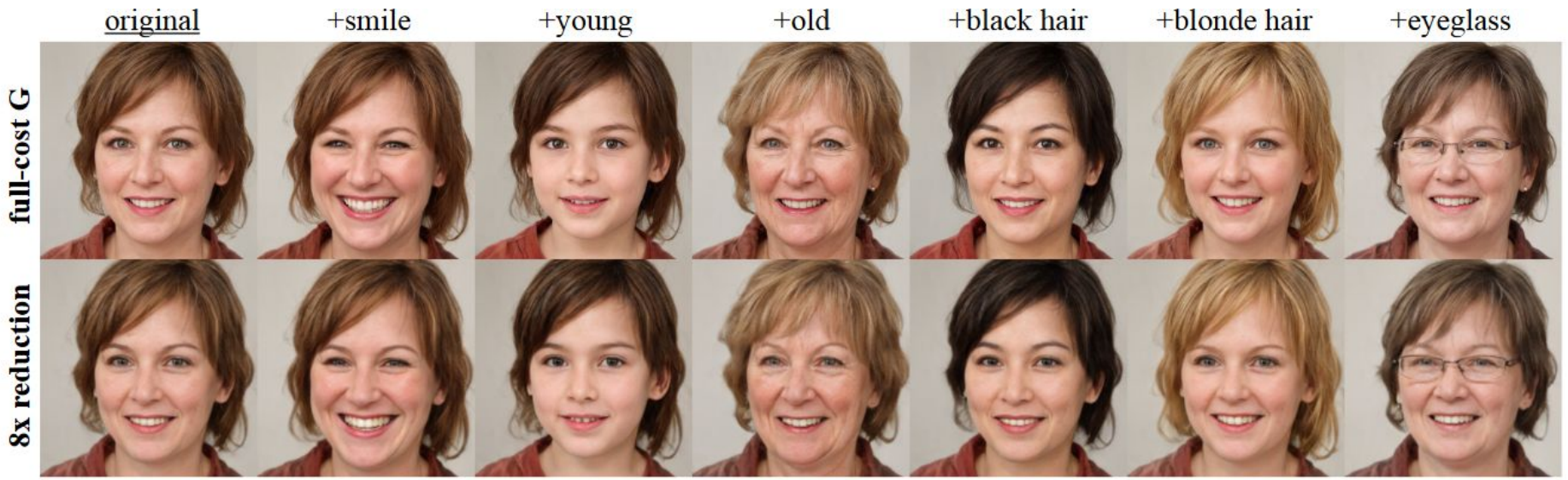
Эксперименты

(b) LSUN Car generated samples

Эксперименты

		FID-50k↓				Path length↓	
Channels:		1.0×	0.75×	0.5×	0.25×	1.0×	0.5×
FFHQ 1024	StyleGAN2	2.84	-	3.31	-	145.0	124.5
	Anycost	2.77	3.05	3.28	5.01	144.2	147.2
Car 512	StyleGAN2	2.32	-	3.19	-	415.5	471.2
	Anycost	2.38	2.46	2.61	3.69	380.1	430.0

Эксперименты



Эксперименты

Table 5: Inference FPS & speed up rates on different devices.

FLOPs reduction	1×	2×	4×	6×	8×	10×
Xeon CPU FPS	0.63	1.78	5.94	6.24	7.48	7.35
speed up rate	1×	2.8×	9.5×	10.0×	11.9×	11.7×
Nano GPU FPS	0.65	1.17	2.77	3.59	4.07	4.1
speed up rate	1×	1.8×	4.2×	5.5×	6.2×	6.3×

Источники

- Anycost GANs for Interactive Image Synthesis and Editing
  Ji Lin, Richard Zhang, Frieder Ganz, Song Han, Jun-Yan Zhu
  https://arxiv.org/abs/2103.03243