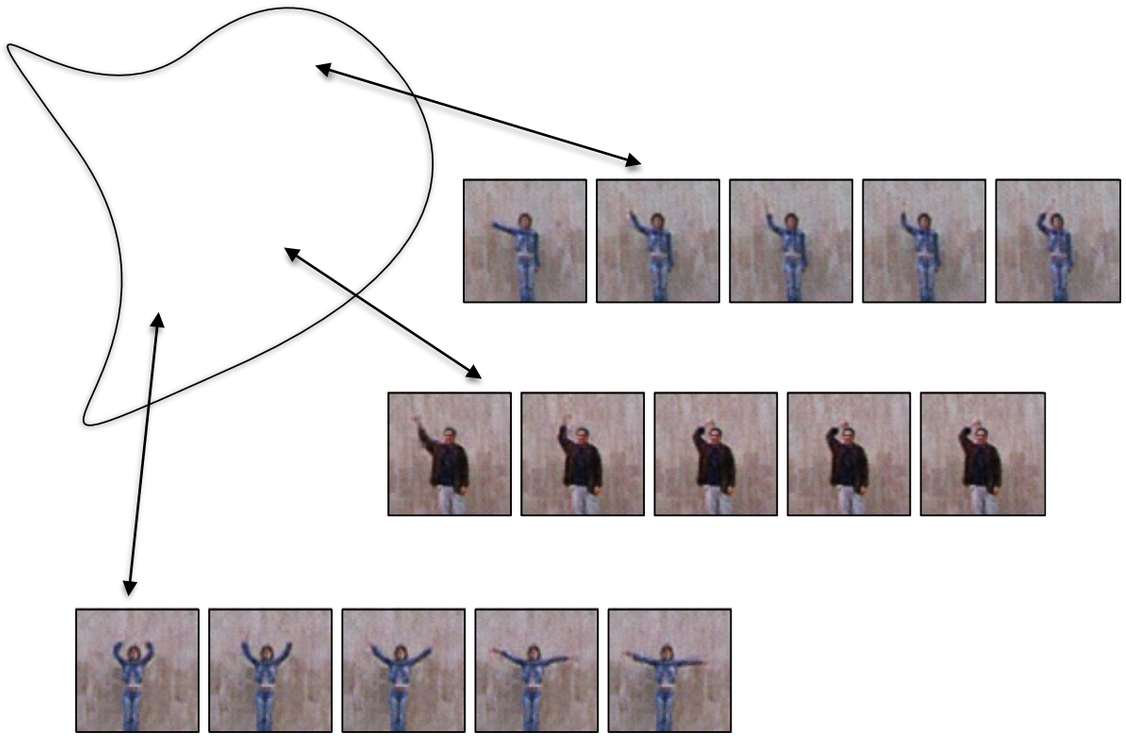
Challenge: Improving MoCoGAN

Sergey Tulyakov

Video representation

Hidden space: every point is an **video**

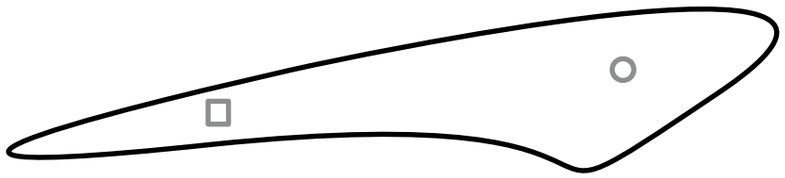


Limitations:

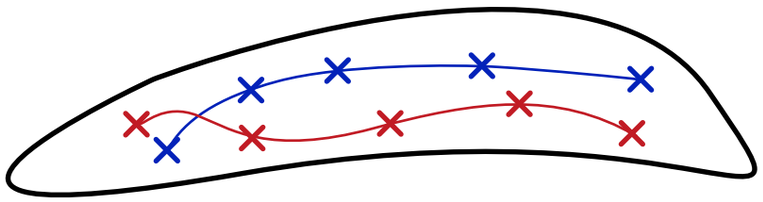
- Fixed length videos only
- No control over motion and content

MoCoGAN representation

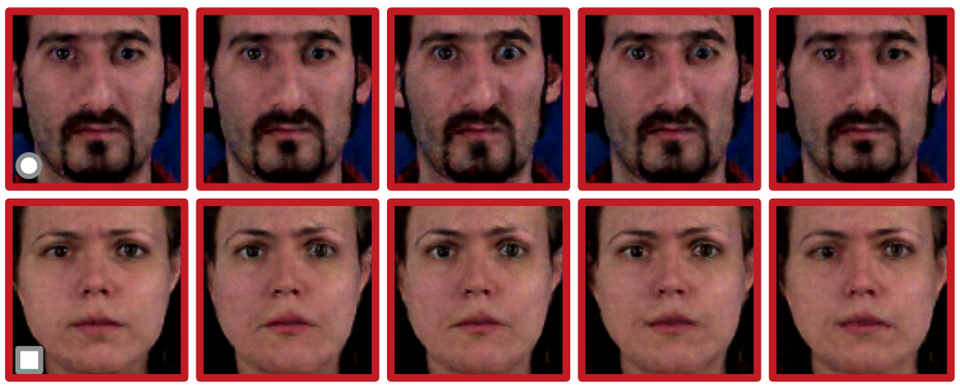
Content subspace



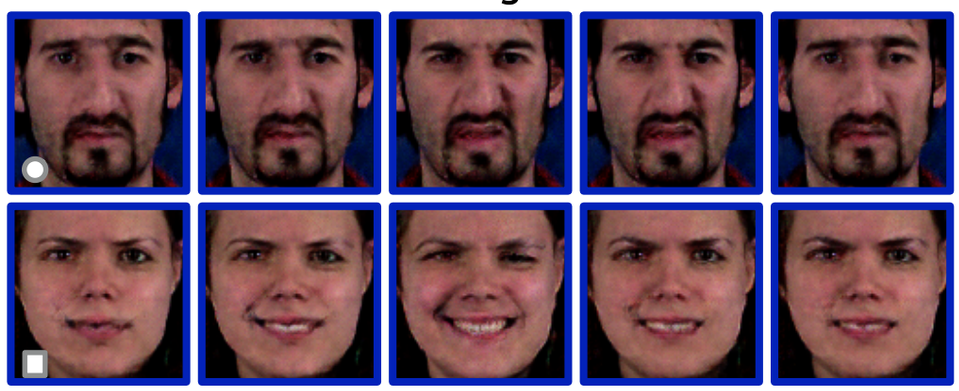
Motion subspace



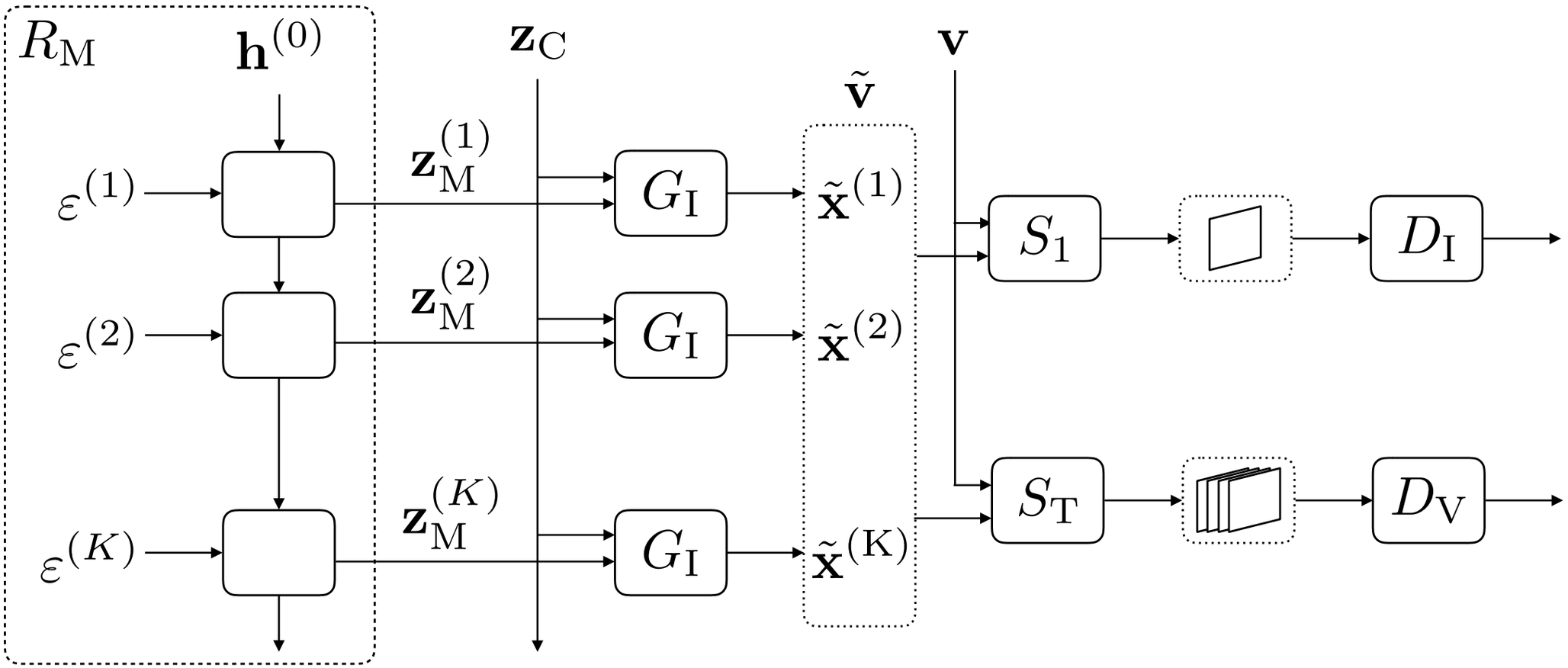
Fear



Disgust



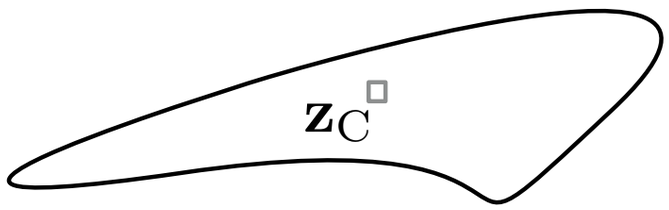
MoCoGAN framework



MoCoGAN framework

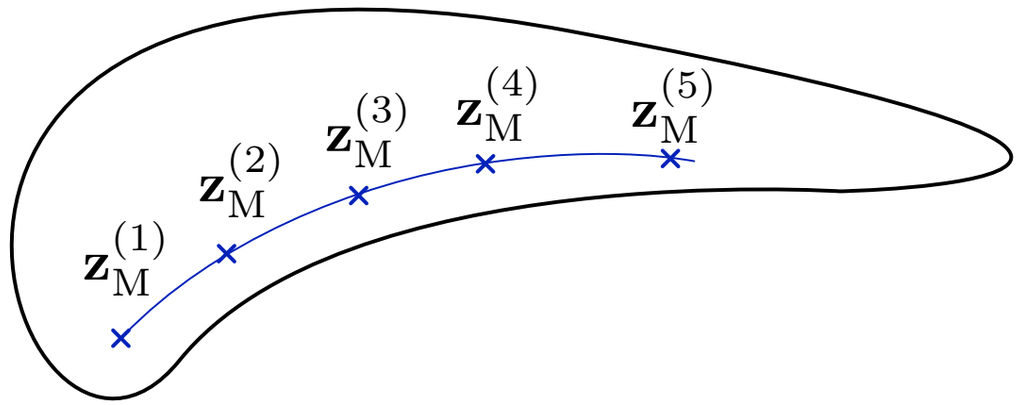
Sampled content

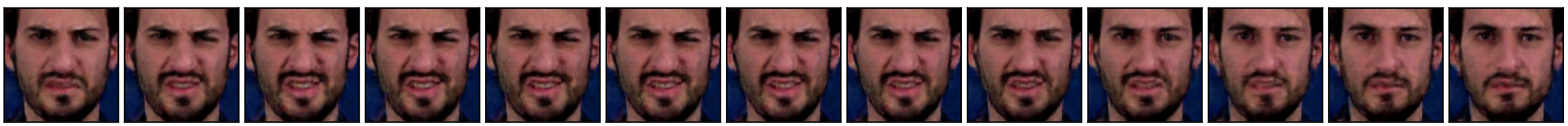
$$\mathbf{Z}_\mathrm{C} = [\mathbf{z}_\mathrm{C}, \mathbf{z}_\mathrm{C}, ..., \mathbf{z}_\mathrm{C}]$$



Motion trajectory

$$\mathbf{Z}_\mathrm{M} = [\mathbf{z}_\mathrm{M}^{(1)}, \mathbf{z}_\mathrm{M}^{(2)}, ..., \mathbf{z}_\mathrm{M}^{(K)}]$$

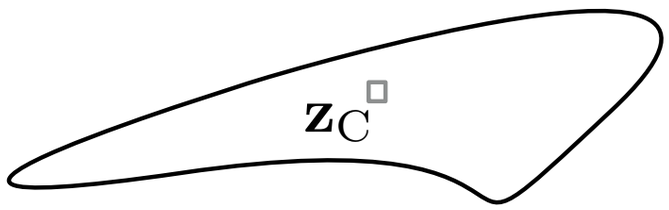




MoCoGAN framework

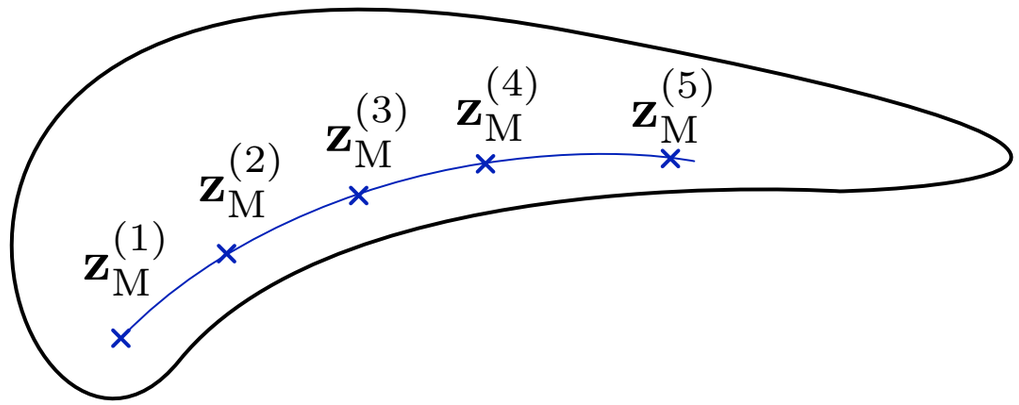
Sampled content

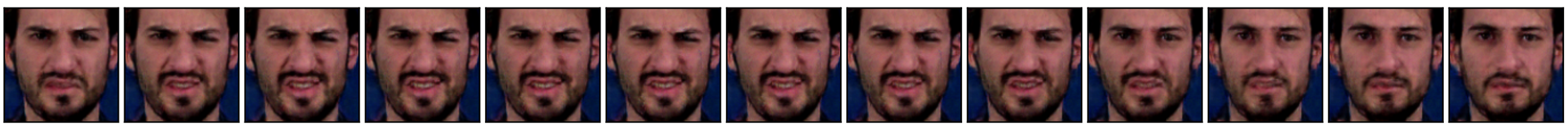
$$\mathbf{Z}_\mathrm{C} = [\mathbf{z}_\mathrm{C}, \mathbf{z}_\mathrm{C}, ..., \mathbf{z}_\mathrm{C}]$$



Motion trajectory

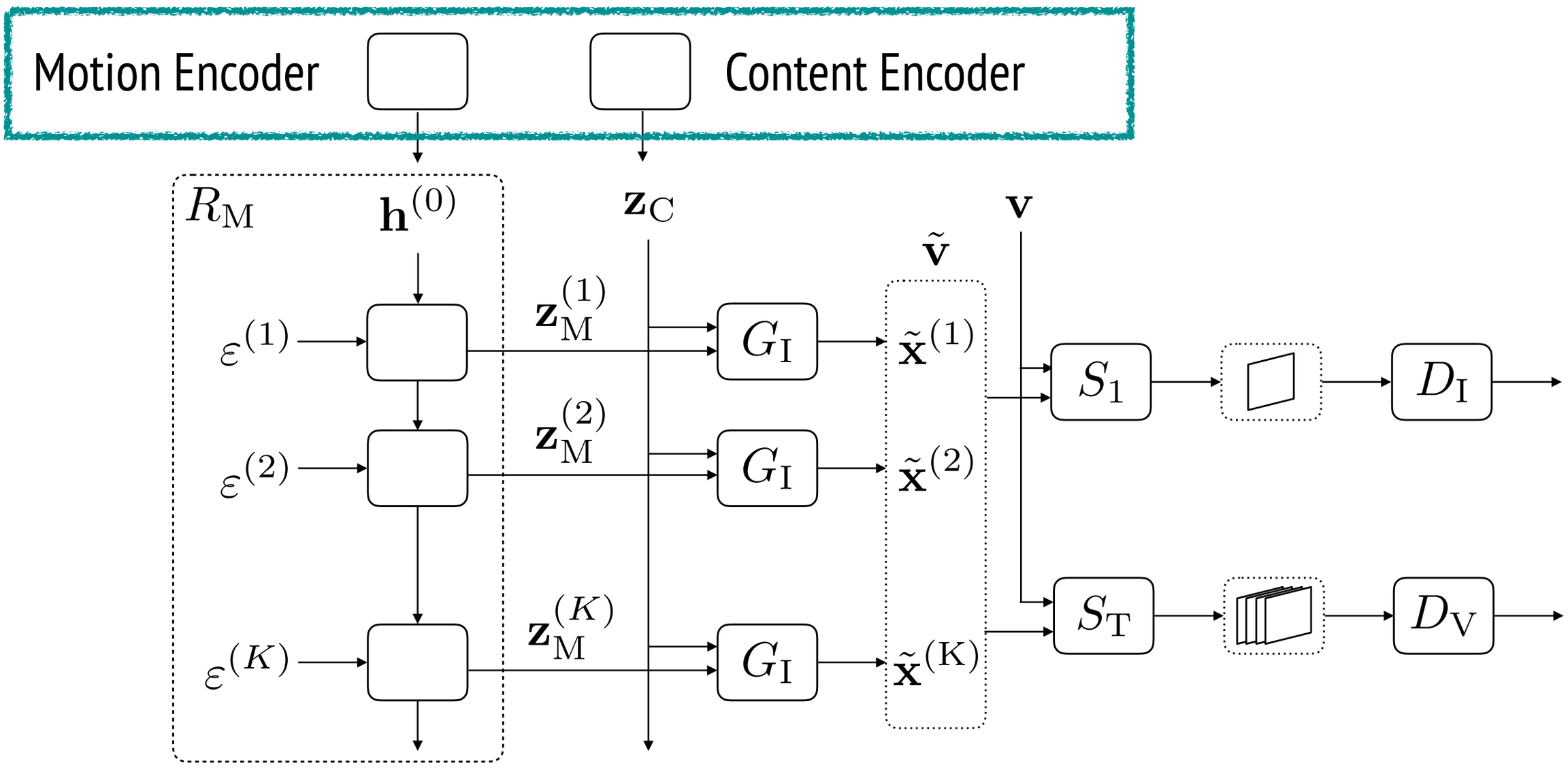
$$\mathbf{Z}_\mathrm{M} = [\mathbf{z}_\mathrm{M}^{(1)}, \mathbf{z}_\mathrm{M}^{(2)}, ..., \mathbf{z}_\mathrm{M}^{(K)}]$$





MoCoGAN for Video Prediction

Your task



MoCoGAN for Video Prediction

Code: https://github.com/sergeytulyakov/mocogan

Paper: https://arxiv.org/abs/1707.04993

Task: Create a Version of MoCoGAN for Video Prediction. Improve quality

Ideas:

- Create a content and motion encoders for 1,2,3 images
- Use CycleGAN-generator to get higher quality of images
- Use motion encoders for the whole video
- Use better data (VoxCeleb for example)

Places to look at:

- sample_z_m: https://github.com/sergeytulyakov/mocogan/blob/master/src/models.py#L217-L229
- sample_z_content: https://github.com/sergeytulyakov/mocogan/blob/master/src/models.py#L249-L257