Graphics Seminar Series

Title - Animating Arbitrary Objects via Deep Motion Transfer

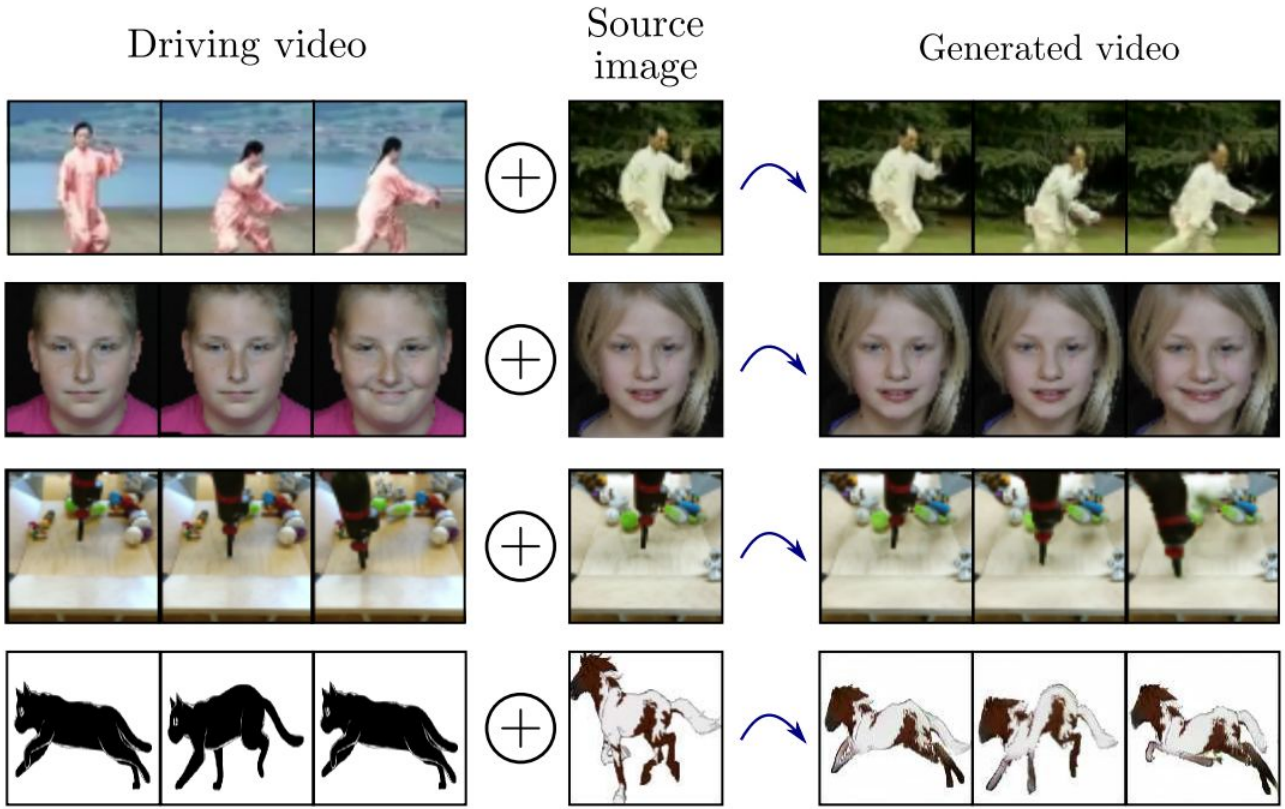
Author's - Aliaksandr Siarohin, Sergey Tulyakov , Elisa Ricci and Nicu Sebe

Venue - CVPR 2019

Presented by - Gaurav Rai

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY **DELHI**

Idea



source - http://www.stulyakov.com/papers/monkey-net.html

Major Contribution of the Paper

- A Keypoint Detector unsupervisely trained to extract object keypoints.
- A Dense Motion prediction network for generating dense heatmaps from sparse keypoints, in order to better encode motion information.
- A Motion Transfer Network, which uses the motion heatmaps and appearance information extracted from the input image to synthesize the output frame.

Problem

- Over the past few years, researchers have developed approaches for automatic synthesis and enhancement of visual data.

- Recent research studies have started exploring the use of deep generative models for image animation and video retargeting.

- However, these approaches have limitations: for example, they rely on pre-trained models for extracting object representations that require costly ground-truth data annotations.

- Furthermore, these works do not address the problem of animating arbitrary objects: instead, considering a single object category or learning to translate videos from one specific domain to another.

Deep Video Generation

- Video generation is closely related to the future frame prediction problem. Given a video sequence, the aim is to synthesize a sequence of images which represents a coherent continuation of the given video.

- Recent video generation approaches used recurrent neural networks within an adversarial training framework.

- The more challenging task: image animation requires decoupling and modeling motion and content information, as well as a recombining them.

Object Animation

- The problems of image animation and video retargeting have attracted attention from many researchers in the fields of computer vision, computer graphics and multimedia.

- Traditional approaches are designed for specific domains, as they operate only on faces, human silhouettes, etc.

- Image animation from a driving video can be interpreted as the problem of transferring motion information from one domain to another.

- The past approaches only learns the association between two specific domains, while we want to animate an image depicting one object without knowing at training time which object will be used in the driving video.

Dataset

They demonstrate the effectiveness of the framework by conducting an extensive experimental evaluation on three publicly available datasets, previously used for video generation:

- The Tai-Chi

  [https://github.com/sergeytulyakov/mocogan]

- The BAIR robot pushing

  [https://www.aminer.org/pub/5a260c2e17c44a4ba8a23f1c/self-supervised-visual-planning-with-temporal-skip-connections]

- The UvA-NEMO Smile datasets

  [https://ivi.fnwi.uva.nl/isis/publications/bibtexbrowser.php?key=DibekliogluECCV2012&bib=all.bib]

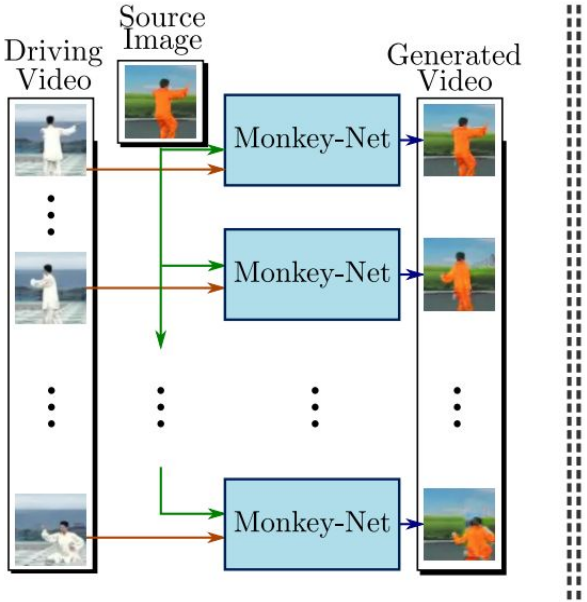
Approach

- This paper addresses some of the limitations by introducing a novel deep learning framework for animating a static image using a driving video.

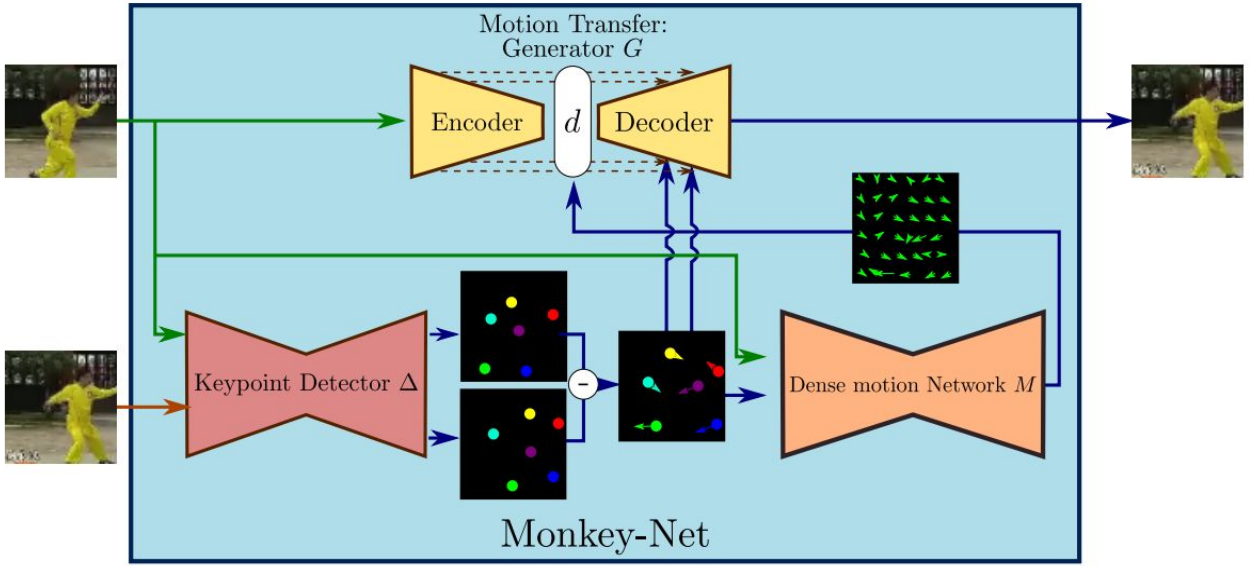
- The framework is not designed for specific object category, but rather is effective in animating arbitrary objects.

- It introduce a novel strategy to model and transfer motion information, using a set of sparse motion-specific keypoints that were learned in an unsupervised way to describe relative pixel movements.

Monkey-Net Architecture



(a) Image Animation



(b) Monkey-Net Architecture for Self-Learned Animation

Experiment

- Evaluating the results of image animation methods is a difficult task, since ground truth animations are not available.

- X2Face is the only previous approach for data-driven model-free image animation.

	Tai-Chi			Nemo			Bair
	$\mathcal{L}_1$	(AKD, MKR)	AED	$\mathcal{L}_1$	AKD	AED	$\mathcal{L}_1$
X2Face	0.068	(4.50, 35.7%)	0.27	0.022	0.47	0.140	0.069
Ours	**0.050**	**(2.53, 17.4%)**	**0.21**	**0.017**	**0.37**	**0.072**	**0.025**

Results

Reference

- http://www.stulyakov.com/papers/monkey-net.html
- https://github.com/AliaksandrSiarohin/monkey-net
- https://aliaksandrsiarohin.github.io/first-order-model-website/


Thank
You