

CarvingNet + Improved Seam Carving for Video Retargeting

20192765 김해우

프로젝트 제안서와의 차이

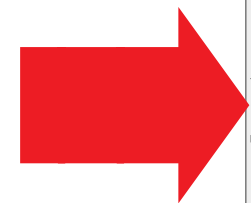
프로젝트 제안서에서 제안한 방식

01 OpenCV를 통해 영상 데이터를 프레임단위로 분할해 이미지 데이터 추출

02 Seam Carving을 통해 과거 영상을 현대의 Display에 맞게 resizing

03 Super Resolution을 통해 초고해상도 이미지로 변환

프레임단위로 추출해 seam carving을 진행할 시 프레임마다 seam의 중요도가 다르게 계산되어 다시 영상으로 전환해 재생 시 시간축으로 연결성을 갖지 못하게 되어 불규칙적으로 움직이는 떨림 현상(jitter)이 발생할 가능성이 높아지게 되어 부적절 하다고 판단.



Seam Carving을 동영상에 적용한 Video Retargeting을 사용하는것이 옳다고 판단되어, 1번과정을 생략하고 2번 3번과정을 image데이터가아닌 video데이터를 통해 진행하는 방법으로 진행

01. CarvingNet

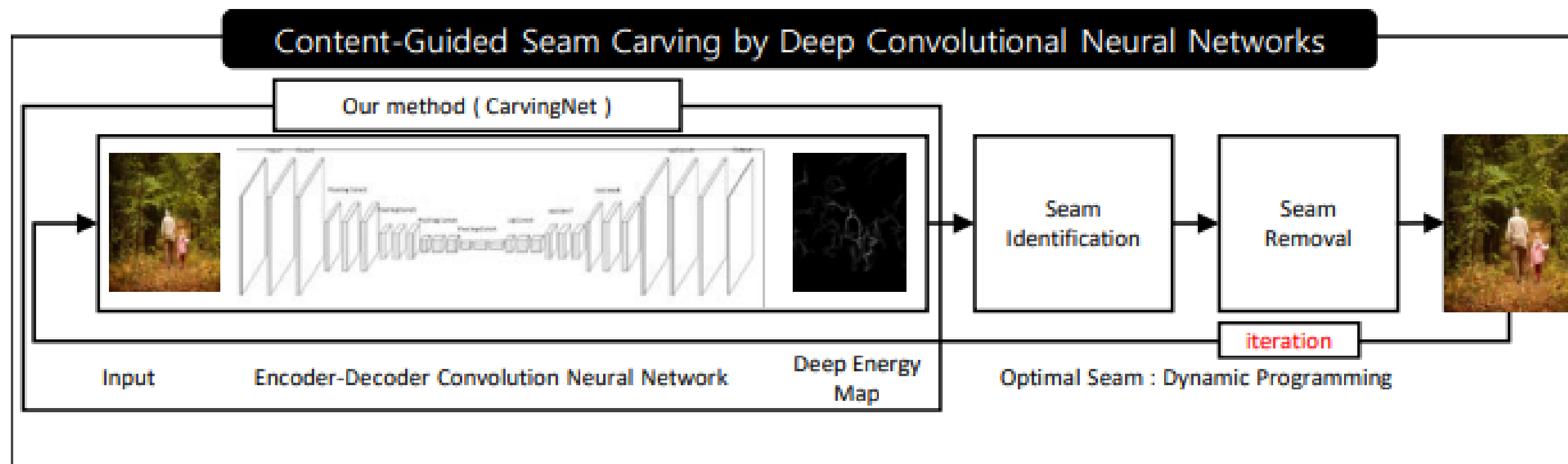
CarvingNet

seam carving의 편향이 발생할때 임계값이 발생하고, 그 값에 의해 energy map의 강도가 다양해진다. 하지만 임계값을 변화시키는 것으로, 다양한 이미지에대하여 일반화해 적용하기 힘들다.

➡ encoder-decoder convolution neural network(E-D CNN)을사용해 다양한 이미지들이 학습된 모델로 energy map을 안정적으로 생성하는 방법을 제시

본 논문에선

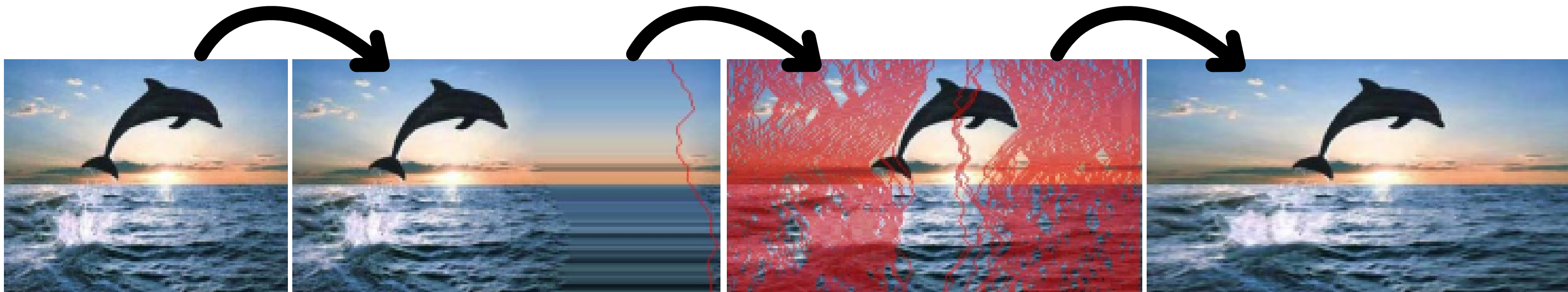
- image retargeting을 seam carving을 통해소개
- DCNN을 사용한 encoder-decoder 구조를 설명하고 DCNN으로 deep energy map을 생성하는 방법을 설명
- 실험환경, dataset, 평가방법, 실험결과 설명



Seam Carving

Seam Carving 기술은 내용을 유지하면서 이미지의 크기를 변환하는 기술
진행 과정으로 간단하게

1. 각 픽셀의 에너지를 계산 : 여러가지 계산 알고리즘이 존재
 2. 계산한 값으로 Seam들의 energy map을 작성 : 값이 낮을수록 중요도가 낮음
 3. 필요한 해상도 만큼 Seam들을 삭제 또는 생성해 input받은 데이터 변환 : 중요도가 낮은 Seam순으로 진행
- 어떤 픽셀이 주위 픽셀과 얼마나 비슷한지 측정하는 에너지 함수를 정의한다. (논문에서는 간단하게 $e = (x\text{방향 픽셀 색깔 변화율}) - y\text{방향 픽셀 색깔 변화율}$) 로 정의해서 사용, e 값이 작을 수록 주위 픽셀과 비슷하다는 뜻)
 - 에너지 합이 가장 작은 1픽셀 너비의 가로줄 혹은 세로줄을 찾는다. (아래 빨간줄 중 한줄)
 - 빨간줄을 제거 또는 복사해 한줄 더 붙임
 - 이 과정을 원하는 크기가 만들어질 때까지 반복

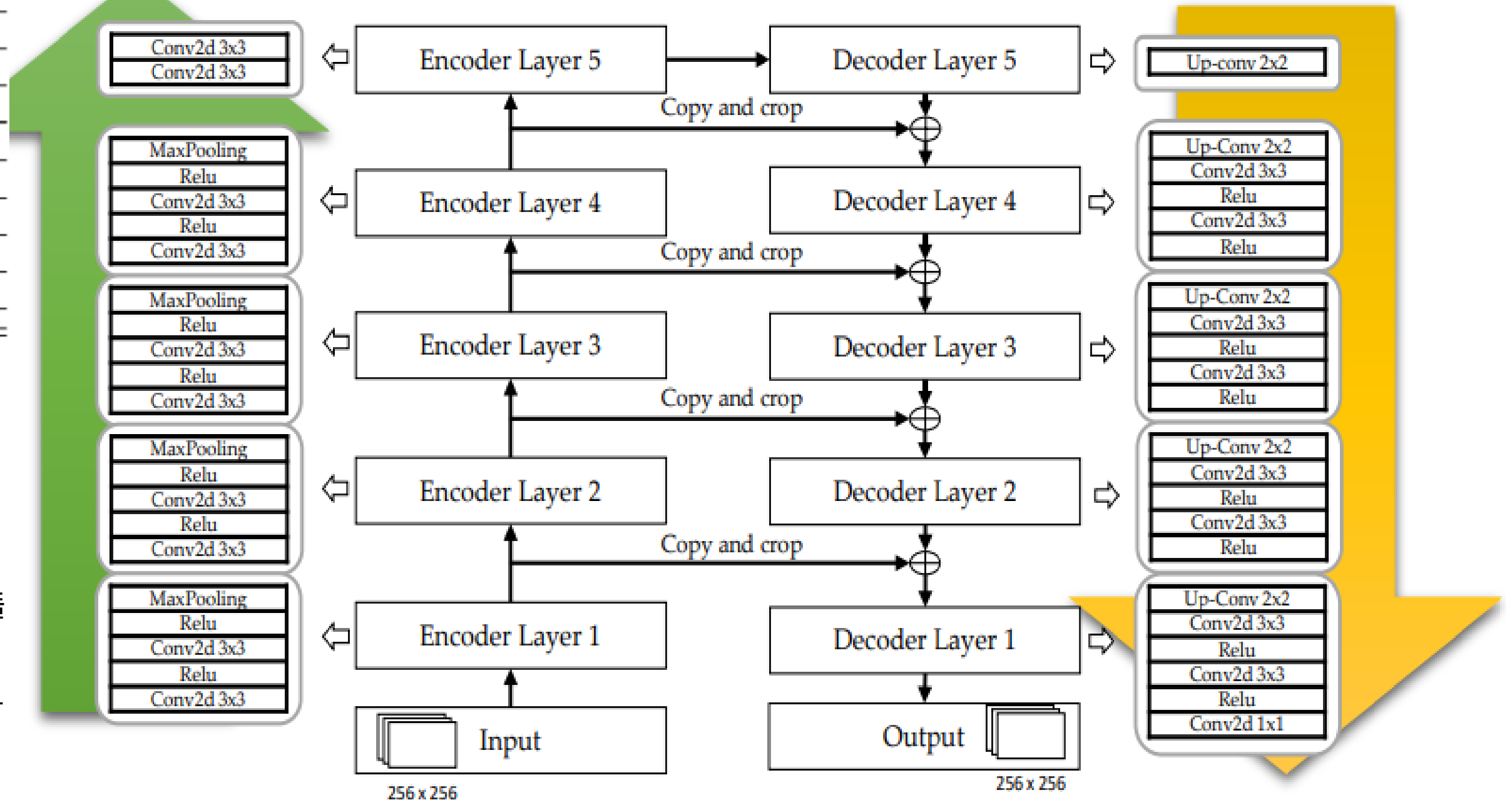


Network Architecture

	Unit level	Conv layer	Filter	Stride	Output size
Input					$256 \times 256 \times 3$
Encoding	Level 1	Conv 1	$3 \times 3/32$	1	$256 \times 256 \times 32$
		Conv 1	$3 \times 3/32$	1	$256 \times 256 \times 32$
	Level 2	Conv 2	$3 \times 3/64$	2	$128 \times 128 \times 64$
		Conv 2	$3 \times 3/64$	1	$128 \times 128 \times 64$
	Level 3	Conv 3	$3 \times 3/128$	2	$64 \times 64 \times 128$
		Conv 3	$3 \times 3/128$	1	$64 \times 64 \times 128$
	Level 4	Conv 4	$3 \times 3/256$	2	$32 \times 32 \times 256$
		Conv 4	$3 \times 3/256$	1	$32 \times 32 \times 256$
Bridge	Level 5	Conv 5	$3 \times 3/512$	2	$16 \times 16 \times 512$
		Conv 5	$3 \times 3/512$	1	$16 \times 16 \times 512$
Decoding	Level 6	De-conv 6	$3 \times 3/256$	1	$32 \times 32 \times 256$
		De-Conv 6	$3 \times 3/256$	2	$32 \times 32 \times 256$
	Level 7	De-conv 7	$3 \times 3/128$	1	$64 \times 64 \times 128$
		De-conv 7	$3 \times 3/128$	2	$64 \times 64 \times 128$
	Level 8	De-conv 8	$3 \times 3/64$	1	$128 \times 128 \times 64$
		De-conv 8	$3 \times 3/64$	2	$128 \times 128 \times 64$
	Level 9	De-conv 9	$3 \times 3/32$	1	$256 \times 256 \times 32$
		De-conv 9	$3 \times 3/32$	2	$256 \times 256 \times 32$
Output		De-conv 10	1×1	1	$256 \times 256 \times 1$

- input/output data는 256x256
- 초록색이 Encoder
- 2x2 Maxpooling으로 이미지 사이즈를 반으로 줄임
- 노란색이 Decoder
- Encoder의 역함수를 적용해 원래대로 사이즈를 되돌림
- pixel의 예측값이 seam이나 최소경로를 반영하게 하기 위해 마지막 layer에 활성화 함수를 sigmoid로 설정

Encoder-Decoder convolution neural network



Generation of energy map using E-D CNN MODEL

- 두 확률분포의 차를 구하기 위해 Cross entropy 사용
- 마지막 cross entropy loss함수가 마지막 feature map과 Sigmoid함수를 합쳐서 energy map의 값을 출력함

$$C = -\frac{1}{n} \sum_x [y \ln f(s_1) + (1-y) \ln (1-f(s_1))] \quad \text{Cost function}$$

$$L = -[y \ln f(s_1) + (1-y) \ln (1-f(s_1))] \quad \text{Loss function}$$

최종 엔트로피양

세로 방향 Seam

가로 방향 Seam

$$\frac{\partial C}{\partial w_j} = \frac{1}{n} \sum_x x_j (\sigma(z) - y)$$

$$S^x = \{S_i^x\}_{i=1}^n = \{(x(i), i)\}_{i=1}^n, \quad \text{where } \forall i, |x(i) - x(i-1)| \leq 1$$

$$S^y = \{S_j^y\}_{j=1}^m = \{(y(j), j)\}_{j=1}^m, \quad \text{where } \forall j, |y(j) - y(j-1)| \leq 1$$

최종 energy map

$$M(i, j) = e(i, j) + \min(M(i-1, j-1), M(i-1, j), M(i, j-1))$$

- n은 학습데이터 수
- y는 학습데이터에 필요한 output value
- f()는 sigmoid함수

Experimental Results

- 평가지표 : 이미지 간의 기하학적 변화를 측정할때 효과적인 방법인 ARS사용
- 데이터셋 : 37개의 이미지(평가지표를 ARS로 뒤 각 픽셀마다를 데이터라 생각하면 돼서 적은양이 아님)

		SC[1]	SCL	SM	WARP[24]	Our Method
ARS	Average	0.909	0.902	0.880	0.908	0.913
	Std	0.053	0.074	0.101	0.064	0.047

다른 Retargeting 기법보다 더 높은 성능을 보임



Original



Scale



Seam Carving



Our

- seam carving (SC)
- scaling (SCL)
- shift map (SM)
- non-homogeneous warping (WARP)

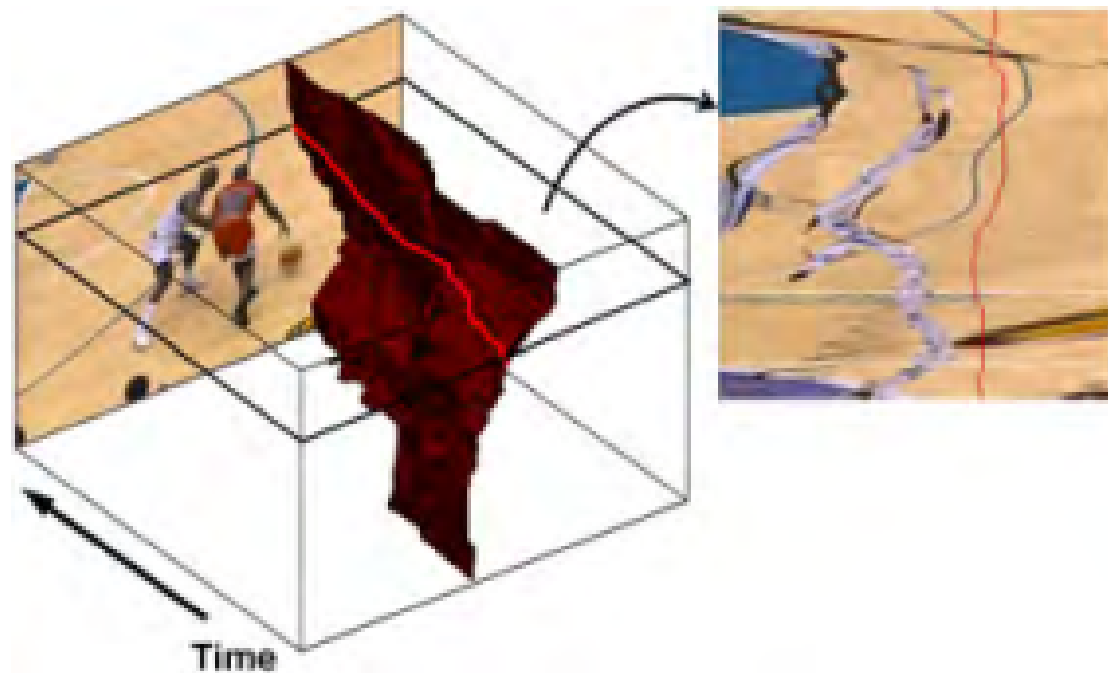
02 ImprovedSeam Carving for

- Video Retargeting

Improved Seam Carving for Video Retargeting

프레임단위로 추출해 seam carving을 토대로한 Retargeting을 진행할 시 프레임마다 seam의 중요도가 다르게 계산되어 다시 영상으로 전환해 재생 시 시간축으로 연결성을 갖지 못하게 되어 불규칙적으로 움직이는 떨림 현상(jitter)이 발생할 가능성이 높아지게 된다.

이를 해결하기 위해 제시된 새로운 방법이 이 논문의 내용



간단히 설명하면 화면(2차원)에 시간의 축을 더해 3차원의 데이터를 가지고 원래 선(1차원)으로 표현되던 Seam을 2(차원)으로 계산해 삭제 또는 더하는 과정을 진행하는 방법론과 진행 방도를 작성한 논문인데 시간이 부족하여 자세한 설명은 이후에 진행하고자 한다 (누락될수도있음)

참고논문

- CarvingNet: Content-Guided Seam Carving using Deep Convolution Neural Network
- Improved Seam Carving for Video Retargeting
- Seam Carving for Content-Aware Image Resizin
- 비디오 리타겟팅(retargeting) 기술 동향
- 고속컨텐츠인식동영상리타겟팅기법

THANK
YOU