

고속 콘텐츠 인식 동영상 리타겟팅 기법

박대현*, 김윤*

Fast Content-Aware Video Retargeting Algorithm

Dae-Hyun Park*, Yoon Kim*

요약

본 논문에서는 동영상의 주요 콘텐츠를 보존하면서 영상의 크기를 변환하는 고속 동영상 리타겟팅 기법을 제안한다. 기존의 Seam Carving에서는 seam을 하나씩 구할 때마다 누적 에너지의 갱신이 발생하며, 여기서 누적 에너지는 동적계획법을 이용하여 계산하기 때문에 전체 연산시간의 지연은 불가피하다. 본 논문에서는 전체 동영상을 특징이 서로 비슷한 scene으로 나누고, 각 scene의 첫 프레임에서는 seam이 될 수 있는 모든 후보들 중 복수개의 seam을 추출하여 누적 에너지의 갱신과정을 줄여 고속화한다. 또한 scene의 두 번째 프레임부터 인접한 프레임 상호간에 상관성을 이용하여, 연속하는 프레임은 누적 에너지를 계산하지 않고 이전 프레임의 seam 정보를 참조한 계산만으로 모든 seam을 추출한다. 따라서 제안하는 시스템은 누적 에너지에 계산되는 연산량을 대폭 줄였으며 전체 프레임의 분석도 필요하지 않아 고속화가 가능하고, 콘텐츠의 떨림 현상은 발생하지 않는다. 실험 결과는 제안하는 방법이 처리 속도와 메모리 사용량 면에서 실시간 처리에 적합하고, 영상이 가지고 있는 콘텐츠를 보존하면서 영상의 크기를 조절할 수 있음을 보여준다.

▶ Keywords : 콘텐츠 인식 비디오 리타겟팅, 심 카빙, 실시간 영상 처리

Abstract

In this paper, we propose a fast video retargeting method which preserves the contents of a video and converts the image size. Since the conventional Seam Carving which is the well-known content-aware image retargeting technique uses the dynamic programming method, the repetitive update procedure of the accumulation energy is absolutely needed to obtain seam. The energy update procedure cannot avoid the processing time delay because of many operations by the image full-searching. By applying the proposed method, frames which have similar features in video are

•제1저자 : 박대현 •교신저자 : 김윤

•투고일 : 2013. 8. 29, 심사일 : 2013. 9. 13, 게재확정일 : 2013. 10. 7.

* 강원대학교 컴퓨터정보통신공학과(Dept. of Computer and Communications Engineering, Kangwon National University)

classified into a scene, and the first frame of a scene is resized by the modified Seam Carving where multiple seams are extracted from candidate seams to reduce the repetitive update procedure. After resizing the first frame of a scene, all continuous frames of the same scene are resized with reference to the seam information stored in the previous frame without the calculation of the accumulation energy. Therefore, although the fast processing is possible with reducing complexity and without analyzing all frames of scene, the quality of an image can be analogously maintained with an existing method. The experimental results show that the proposed method can preserve the contents of an image and can be practically applied to retarget the image on real time.

▶ Keywords : Content-Aware Video Retargeting. Seam Carving. Real-Time Image Processing

I. 서 론

유무선 통신의 발전으로 휴대용 단말에 대한 사용이 급증하였고, 이러한 수요에 편승하여 많은 종류의 휴대용 단말이 출시되고 있다. 이러한 휴대용 단말은 사용자의 편의를 고려하여 디스플레이 크기와 해상도를 다양하게 생산하며, 각각의 휴대용 단말 디스플레이의 크기에 올바른 출력을 위해서는 영상 크기변환이 필요하다. 하지만 Bilinear^[1]나 Bicubic^[2]과 같은 널리 사용되는 보간법을 사용한 크기변환 기법은 영상의 모든 부분을 동일한 비율로 변경하기 때문에 영상이 가지고 있는 주요 콘텐츠가 줄어들거나 늘어나는 왜곡이 발생한다(그림 1. (b)). 따라서 효과적인 영상 크기 변환을 위해서는 영상이 가지고 있는 콘텐츠를 보존하면서 영상의 크기를 변경하는 콘텐츠 인식 기반의 영상 리타겟팅 기법이 필요하다.

가장 단순한 방법으로 영상내의 주요 콘텐츠를 찾아 해당 콘텐츠만을 잘라서 보여주는 Cropping 기법^{[3][4]}이 있다(그림 1. (c)). 하지만 이 기법은 주요 콘텐츠 이외에 영상 정보의 많은 부분이 삭제된다는 단점이 있다. 이를 개선하여 주요 콘텐츠를 보존하고 전체 영상 정보를 축약하여 보여주는 Fish-Eye Warping 기법^[5]이 있다(그림 1. (d)). 하지만 Fish-Eye Warping 기법도 영상의 주요 콘텐츠를 최대한 보존하는 반면, 영상의 외곽부분으로 갈수록 영상정보들이 심하게 왜곡된다는 단점이 있다. Seam Carving^[6]은 영상 내에서 중요도가 낮은 픽셀들을 삭제 혹은 추가하는 방법으로 다양한 영상 리타겟팅 기법들 가운데서 높은 콘텐츠의 보존 상태를 보여준다(그림 1. (e)). 하지만 알고리즘을 적용하기 위해서는 누적에너지 계산 및 seam을 찾는 과정 등에서 많은

연산량이 필수적이다.

콘텐츠 인식 기반의 영상 리타겟팅을 동영상에 적용하기 위해서는 시간 영역에 대한 고려가 필요하다. 시간축으로 연결성을 갖지 못한다면 영상 내에서 콘텐츠는 불규칙적으로 움직이는 떨림 현상(jitter)이 발생할 수 있다. 떨림 현상이 방지되도록 Seam Carving을 동영상에 적용한 Video Caving^{[7][8]}은 시간축으로 프레임들을 연결한 3차원 큐브(cube)에 그래프 컷(graph cut)을 적용하였다. 하지만 3차원 큐브를 구성하기 위해서는 많은 메모리가 필요하고, seam의 공간과 시간 연결성을 만족하기 위해서는 동영상 전체를 탐색하고 분석하는 작업이 필수적이므로 많은 연산량이 필요하다. Continuous Seam Carving과 선택적인 scaling을 사용하여 결과영상의 안정성을 높인 방법도 제안되었지만, 적용적으로 seam을 추출하는 과정에서 multiscale contrast 기반의 saliency 정보를 사용하여 많은 연산량이 필요한 단점이 있다[10]. Scalable and Coherent Video Resizing은 각 프레임마다 독립적으로 리타겟팅과 warping을 수행하고, optical flow를 통해 전체 프레임의 모션정보를 추출하여 각 프레임의 시간축에 대해 연결성을 구성한다. 하지만 많은 연산량이 필요한 알고리즘이 다수 사용되어 복잡도가 높다[11]. 따라서 모바일 단말과 같은 한정된 자원을 가진 시스템에서 연속적으로 입력되는 영상들에 대해 실시간으로 처리할 수 없다.

본 논문에서는 고속의 콘텐츠 인식 동영상 리타겟팅 기법을 제안한다. 먼저, 동영상상에서 효과적으로 scene을 분리하기 위하여, 명도 값의 차이와 히스토그램 값의 차이를 이용하여 장면전환(scene change)을 검출한다. 장면전환이 발생하여 새로운 scene이 시작되면, scene의 첫 프레임에서는 개선된 다중 seam 추출 Seam Carving 기법을 이용하여 영상

의 크기변환을 한다. 기존의 Seam Carving 방법은 반복적인 누적 에너지의 갱신 연산과정으로 한 번에 하나의 seam만을 추출하는 단점이 있다. 따라서 여러 개의 seam을 추출하기 위해서는 seam의 개수만큼 반복적인 연산이 필요하다. 이러한 반복적인 연산과정을 줄여 처리 속도를 개선하기 위해 후보 seam들의 공간의존도를 계산한다. 공간의존도를 이용하여 다수의 seam을 추출함으로써 갱신과정의 반복 횟수를 줄여 처리속도를 향상한다. 공간의존도와 다수의 seam 선택은 좋은 결과영상을 얻기 위한 trade-off가 있기 때문에 적응적으로 threshold를 결정하는 방법을 사용한다. 또한 동일한 scene의 연속하는 프레임들은 이웃하는 프레임과 상관성(correlation)이 크기 때문에 프레임간의 seam들은 유사하게 나타날 가능성이 높다. 이와 같은 상관성을 이용하여 이전 프레임에서 추출된 seam과 seam의 위치에 따른 energy 정보를 참조하여 현재 프레임의 새로운 seam을 계산한다. 여기서 energy정보로 temporal forward energy를 제안한다. 본 논문에서 제안하는 시스템은 연산량을 대폭 줄이고 3차원 큐브를 구성하여 전체 프레임 영상에 대해 탐색 및 분석을 필요로 하지 않기 때문에 프레임 단위로 고속 처리가 가능하다.

본 논문의 구성은 다음과 같다. 2장에서 Seam Carving 방법에 대한 전체적인 구성과 문제점에 대하여 기술하고, 3장에서는 제안하는 알고리즘을 소개하며, 4장에서는 제안한 알고리즘을 기존의 방법들과 비교하는 실험을 통하여 결과의 우수성을 입증한다. 최종적으로 5장에서 결론을 맺는다.

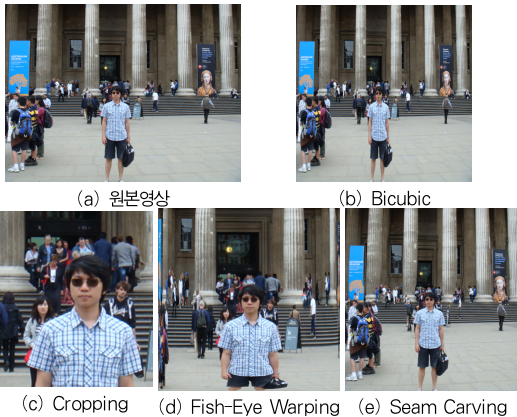


그림 1. 다양한 영상 크기 변환 방법 비교
Fig. 1. Comparison of various image resizing method

II. Seam Carving

Seam Carving은 영상 내에서 에너지 변화가 가장 낮은 seam을 추출하여 영상의 크기를 조절한다. seam은 $W \times H$ 크기의 영상에서 식 (1)과 같이 정의된다.

$$\begin{aligned}
 \mathbf{s}^v &= \{(\mathbf{X}(j), j)\}_{j=1}^H, \\
 & \text{s.t. } \forall j, |\mathbf{X}(j) - \mathbf{X}(j-1)| \leq 1, \\
 \mathbf{s}^h &= \{(i, \mathbf{Y}(i))\}_{i=1}^W, \\
 & \text{s.t. } \forall i, |\mathbf{Y}(i) - \mathbf{Y}(i-1)| \leq 1, \\
 \mathbf{X} &: \{1, \dots, H\} \rightarrow \{1, \dots, W\}, \\
 \mathbf{Y} &: \{1, \dots, W\} \rightarrow \{1, \dots, H\}.
 \end{aligned} \tag{1}$$

\mathbf{s}^v 는 세로 방향의 seam, \mathbf{s}^h 는 가로 방향의 seam을 나타내고, \mathbf{X} 와 \mathbf{Y} 는 맵핑(mapping)을 나타내며 영상의 행 좌표와 열 좌표를 나타낸다. 즉, 세로방향의 seam은 동일한 행에 하나의 픽셀만을 가지며 세로로 연결된 좌표 집합이고, 가로방향의 seam은 동일한 열에 하나의 픽셀만을 가지며 가로로 연결된 좌표 집합이다.

영상에서는 다수의 seam들을 추출할 수 있으며, 그 중에 얻고자 하는 최적의 seam \mathbf{S}^* 은 식 (2)와 같이 정의한다.

$$\mathbf{S}^* = \min_{a \in \mathcal{S}} \{E(a)\}. \tag{2}$$

\mathcal{S} 는 영상에서 얻을 수 있는 모든 seam의 집합이며, $E(\cdot)$ 는 seam에 대한 누적 에너지 함수이다. 각 seam들의 $E(\cdot)$ 중에 최소값을 구하기 위해서는 누적 에너지 M 이 필요하며, M 은 동적계획법을 통해 값을 도출할 수 있다. 식 (3)은 $W \times H$ 크기의 영상에서 세로 방향의 seam 구하는 방법을 나타낸다.

$$\begin{aligned}
 M(i, j) &= e(i, j) + \min \left\{ \begin{aligned} &M(i-1, j-1), \\ &M(i, j-1), \\ &M(i+1, j-1) \end{aligned} \right\}, \\
 0 &\leq i < W, 0 \leq j < H.
 \end{aligned} \tag{3}$$

여기서 $e(\cdot)$ 는 영상에서 해당 픽셀위치의 에너지를 구하는 함수이다. 식 (3)을 영상에 적용하여 도출된 M 의 마지막 행에는 세로 방향으로 누적된 에너지 값들이 저장된다(그림 2. (c)). 누적 에너지 영상은 $[0, 255]$ 범위의 gray영상으로 정규화(normalization)하여 나타낸 것으로 밝을수록 에너지 값이 크며, 어두울수록 작은 에너지 값을 가진다. 가로

방향에 대해서도 동일한 방법으로 구할 수 있다(그림 2. (d)). 최적의 seam은 마지막 행의 누적 에너지 값 중에 최소 값을 선택하여 역탐색하는 것으로 구할 수 있으며, 그림 2. (b)에서 녹색으로 표시된 픽셀은 가로방향과 세로방향의 최적의 seam을 나타낸 것이다.

Seam을 이용하여 영상의 크기를 조절할 때에는 seam에 해당하는 위치에 픽셀을 추가하거나 삭제하여 크기를 변경한다. 영상의 크기를 한 픽셀이상 조절할 경우에는 복수개의 seam이 필요하다. 복수개의 seam을 추출하기 위해서는 이미 추출한 seam을 구성하는 픽셀들을 영상에서 제외 한 후에 다시 M 의 갱신을 통하여 식 (1)을 만족하는 다음 순번의 seam을 추출한다. 여기서 M 의 갱신으로 인한 전체 처리 시간 지연은 불가피하다. 만약 입력영상과 크기를 변경할 영상의 해상도 차이가 크다면 전체 처리시간은 크게 증가한다.

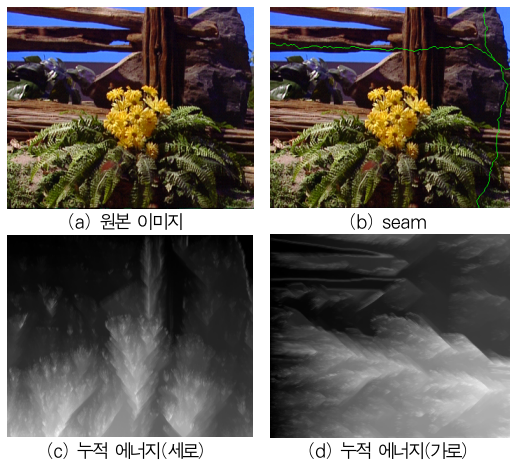


그림 2. 각 방향의 seam과 누적 에너지
Fig. 2. Seam and accumulative energy for each direction

III. 제안하는 알고리즘

본 논문의 고속 동영상 리타캐팅 기법을 위한 시스템은 그림 3과 같이 구성된다. 입력되는 영상에 대해 장면 전환 검출을 통하여 새로운 scene이 시작되었는지 여부를 판단한다. 새로운 scene이 시작되었을 경우에는 입력 프레임에 대해 다중 seam 추출 Seam Carving을 적용하여 seam을 구한다. 이 때 도출된 seam에 대한 정보는 다음 프레임에서 참조될 수 있기 때문에 buffer에 저장하고, seam의 정보를 이용하여 입력 영상을 크기 변환하여 결과영상을 얻는다. 반면에 장면

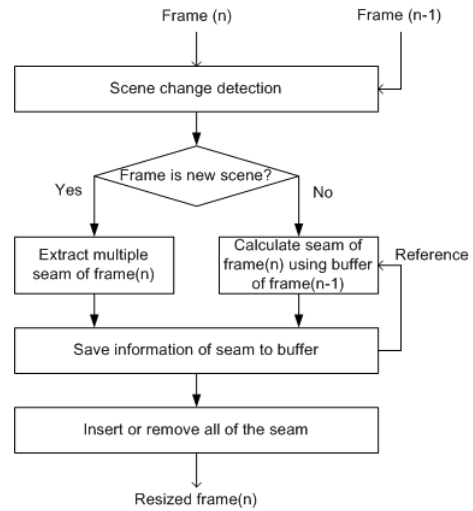


그림 3. 전체 시스템 블록 다이어그램
Fig. 3. Overall system block diagram

전환이 발생하지 않은 경우에는 이전 프레임의 저장된 seam 정보를 이용하여 현재 프레임에서 seam을 계산하여 이를 통해 결과영상을 얻는다.

장면 전환 검출은 명도와 히스토그램을 이용한 특징점을 구해 수행한다. 식 (4)는 $W \times H$ 크기 영상의 이전 프레임 내에서 존재하는 특징 값을 나타낸다.

$$f_{hn} = \sum_{k=0}^{255} |h_n(k) - h_{n-1}(k)|, \tag{4}$$

$$f_{in} = \sum_{j=0}^{H-1} \sum_{i=0}^{W-1} |i_n(i,j) - i_{n-1}(i,j)|.$$

여기서, $h_n(k)$ 는 n 번째 프레임에서 k 의 픽셀값을 가진 히스토그램을 나타내며, 모든 $h(k)$ 의 차이의 합을 f_h 로 정의한다. 또한, $i_n(i,j)$ 는 n 번째 프레임의 (i,j) 번째 위치에서의 픽셀값이며, f_i 는 모든 픽셀의 명도 차이의 합을 나타낸다. 장면 전환은 f_i 와 f_h 중에서 상대적인 차이에 따른 적응적인 임계값보다 값이 크게 나타날 경우에 발생되었다고 판단한다. Scene이 시작되는 첫 번째 프레임에 대해서는 공간의존도와 적응적 임계값을 도입한 방법을 적용하여 복수개의 seam을 추출한다. 기존의 Seam Carving에서는 누적 에너지 M 의 갱신이 구하고자 하는 복수개의 seam 개수와 동일하게 반복적으로 계산되지만, 본 논문에서는 누적 에너지 M 에서 하나 이상의 seam을 획득함으로써 M 의 갱신에 사용되는 반복과정을 줄여 전체 처리 시간을 단축한다. 복수개의 seam을 획

특하기 위해 사용되는 공간의존도는 추출하는 seam들이 식 (1)을 만족하고 공간적으로 독립되도록 만들며, 적응적 임계값은 주요 콘텐츠의 에너지가 높은 픽셀이 seam으로 선택되어 왜곡이 발생하는 것을 방지한다.

복수개의 seam을 추출하기 위해 누적 에너지 M 에서는 누적값이 낮은 순서로 공간의존도를 계산한다. Seam의 공간의존도란 x 방향(세로 방향의 seam) 또는 y 방향(가로 방향의 seam)으로 seam이 생성되어 seam을 구성하는 픽셀 전체 영역을 의미한다.

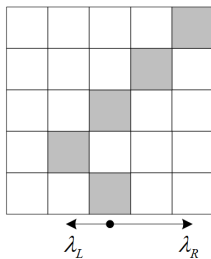


그림 4. Seam의 공간의존도 예시
Fig. 4. Example of spatial dependency of seam

그림 4는 마지막 행에 있는 seam의 좌표를 기준으로 공간의존도 λ_L 와 λ_R 를 나타내며, 이 때 $\lambda_L = 1$, $\lambda_R = 2$ 의 값을 갖는다. 식 (5)는 $W \times H$ 크기의 영상에서 세로 방향의 누적 에너지에 대한 공간의존도를 구하는 식이다. δ_l, δ_r 은 seam의 위치에 따른 각 방향의 공간의존도 변화량이며, 최종 공간의존도인 λ_L, λ_R 은 M 과 동일하게 마지막 행(또는 열)에서 얻어진다. 이렇게 구해진 공간의존도는 seam간의 공간독립성을 만족하지만, 누적 에너지의 값이 큰 seam은 에너지가 높은 픽셀을 포함할 가능성이 크기에 필연적으로 영상 왜곡이 나타날 가능성이 높아진다. 따라서 데이터 그룹의 분산을 이용하는 이진화 알고리즘(9)을 누적 에너지 M 에 적용하여 누적 에너지 값이 작은 부분과 큰 부분으로 분리하는 적응적인 임계값을 구한다. 적응적 임계값보다 작은 누적 에너지 값은 seam이 될 수 있다는 유효성을 판단하는 기준이 된다.

그림 5는 $W \times H$ 크기의 영상에서 적응적 임계값 T 와 공간의존도 λ_L 와 λ_R 를 이용하여 세로방향에 대한 seam의 추출 가능 여부를 확인하는 과정이다. F 는 이미 추출한 seam들이 침범하지 않은 가로 영역으로써 또 다른 seam을 추출할 수 있는 영역이다. 그리고 F 가 포함하는 위치 내에서 가장 작은 누적 에너지 값을 t_c , 그 때의 x 좌표는 c 이다. 최적의 seam은 t_c 의 좌표 c 로부터 역탐색을 통하여 얻어지므로, F 의 범위 내에서 t_c 와 c 는 최적 seam 정보이다. λ_c 는 좌표 c

$$\begin{cases} \text{if } m_e \equiv M(i-1, j-1) \text{ then} \\ \delta_l(i, j) = \delta_l(i-1, j-1) + 1, \\ \delta_r(i, j) = \delta_r(i-1, j-1) - 1 \\ \text{else if } m_e \equiv M(i, j-1) \text{ then} \\ \delta_l(i, j) = \delta_l(i-1, j-1), \\ \delta_r(i, j) = \delta_r(i-1, j-1) \\ \text{else if } m_e \equiv M(i+1, j-1) \text{ then} \\ \delta_l(i, j) = \delta_l(i-1, j-1) - 1, \\ \delta_r(i, j) = \delta_r(i-1, j-1) + 1 \end{cases} \quad (5)$$

$$\begin{aligned} \text{where } m_e &= \min \begin{Bmatrix} M(i-1, j-1), \\ M(i, j-1), \\ M(i+1, j-1) \end{Bmatrix}, \\ \lambda_L &= \min \delta_l, \lambda_R = \max \delta_r, \\ (0 \leq i < W, 0 \leq j < H). \end{aligned}$$

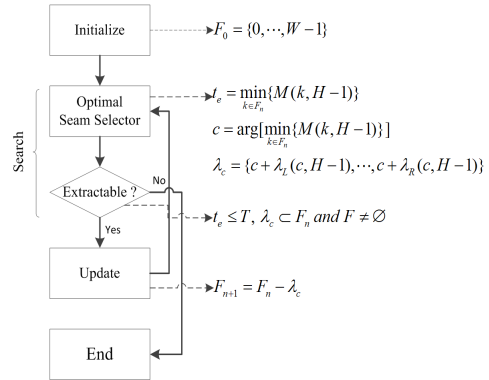


그림 5. Seam 추출 가능 여부 확인 과정
Fig. 5. Confirmation process of possibility of seam extraction

로 시작하는 seam이 포함하는 가로영역이며, 공간의존도 λ_L 과 λ_R 에 의하여 구한다. n 은 한번 계산된 M 내에서 추출한 seam의 개수이다. *Initialize* 과정에서는 F 를 x 좌표($0 \leq x < W$)로 초기화하며, *Search* 과정은 실질적인 seam을 추출하는 부분으로써 F 의 집합영역에 존재하는 x 좌표 영역에 대해서만 탐색하며, seam의 결정은 누적 에너지 값인 t_c 가 적응적 임계값 T 보다 작은 경우에 추출된다. *Update*는 M 에서 추가적인 seam 추출을 위하여 추출한 seam의 λ_c 를 F 에서 제외시키는 과정이다. 이러한 *Search* 과정과 *Update* 과정을 수행하여 한 번 계산된 M 으로부터 하나 이상의 seam을 얻을 수 있다. 한번 계산된 M 에서는 $F \equiv \emptyset$ 이거나 $t_c > T$ 가 될 때까지 추가적인 seam을 얻을 수 있으므로, 누적 에너지 M 의 반복적인 계산을 줄여 고속처리가 가능하다.

입력 프레임에 대해 모든 seam을 구한 후 각각의 seam 정보는 동일한 scene의 다음 프레임에서 참조되기 위해 버퍼

에 저장된다. 식 (6)은 $W \times H$ 크기의 입력영상에서 도출된 seam들이 저장되는 정보를 나타낸다.

$$\begin{aligned}
 S_f &= [S_1, S_2, S_3, \dots], \quad S = (C, E, e_t), \\
 C &= [x_0, x_1, \dots, x_{W-1}] \text{ or } [y_0, y_1, \dots, y_{H-1}], \\
 E &= [e_0, e_1, \dots, e_{W-1}] \text{ or } [e_0, e_1, \dots, e_{H-1}].
 \end{aligned} \quad (6)$$

S_f 는 하나의 프레임에서 구한 seam들의 배열이며, 집합 S 는 하나의 seam에 대한 정보를 나타낸다. C 는 seam의 좌표 배열이고, E 는 각 좌표에서의 에너지를 나타내는 배열이며, e_t 는 E 값을 모두 누적한 값이다.

버퍼에 저장된 seam 정보는 연속적으로 입력되는 프레임의 seam을 추출하는데 참조된다. 프레임마다 추출될 seam은 프레임의 진행방향에 대해 연속성이 존재해야 하며, 이를 위해 같은 scene내에 연속하는 두 프레임들은 서로 유사하다는 상관성을 이용한다. 프레임 간에 상관성이 존재하므로 두 프레임의 에너지도 유사하게 나타나며, 두 영상의 seam도 유사한 위치의 형태로 나타난다.

먼저 현재 프레임에서 seam을 구할 때에 연속하는 프레임 간의 상관성을 이용하여 이전 프레임의 seam 정보를 참조한다. 식 (7)은 이전 프레임에서 저장된 seam정보를 현재 프레임이 참조하는 것을 나타낸다.

$$S_{ref} = S_{f(n-1)}(i). \quad (7)$$

$n-1$ 은 이전 프레임 번호이고, i 는 추출할 seam의 순번이다. 따라서 S_{ref} 는 새로운 seam을 생성하기 위해 참조되는 이전 프레임의 seam이며, 식 (6)의 S 와 같은 데이터 구조를 가지고 있다. 현재 프레임에서 구하고자 하는 seam은 참조 seam인 S_{ref} 의 인접영역에서 형성되며, seam을 구성하는 픽셀의 위치값을 가지는 집합 p 로 형성된다. 시작위치는 p_n 이며 seam을 구성하는 다음 위치의 픽셀인 p_{n+1} 은 공간적 연결성과 시간적 연결성을 동시에 만족해야 한다. 공간적 연결성은 p_n 과 p_{n+1} 이 공간적으로 연결되어야 함을 의미하고, 시간적 연결성은 p_{n+1} 와 $C(n+1) (C_{ref})$ 가 시간적으로 연결되어야 함을 의미한다. 식 (8)은 두 조건을 만족하여 후보 seam의 픽셀로 구성될 수 있는 집합인 $CandidatePr$ 를 구하는 과정이다.

$$\begin{aligned}
 CandidatePr &= SPA \cap TEM, \\
 SPA &= \{s | p_n - 1 \leq s \leq p_n + 1\}, \\
 TEM &= \{t | C \subset S_{ref}, \\
 &\quad C(n+1) - 1 \leq t \leq C(n+1) + 1\}.
 \end{aligned} \quad (8)$$

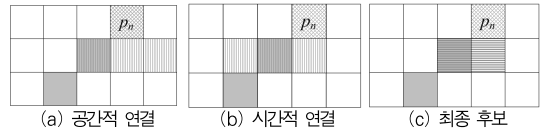


그림 6. 픽셀 후보 예시
Fig. 6. Example of pixel candidate

n 은 p_{n+1} 의 x (가로 방향의 seam) 또는 y (세로 방향의 seam)좌표이며, SPA 와 TEM 은 위의 조건인 공간적 연결과 시간적 연결을 만족하는 픽셀 집합이다. 그림 6은 p_{n+1} 의 위치에서 가능한 SPA , TEM , $CandidatePr$ 를 표현한 예이다. 회색 박스로 표시된 픽셀이 $C (C_{ref})$ 이며, (a)와 (b)에서 세로줄 박스로 표시된 픽셀이 SPA 와 TEM 이다. 첫 번째 줄의 그물모양 박스로 표시된 것은 후보 seam의 시작 위치인 p_n 이다. (c)의 가로줄 영역은 (a)와 (b)를 모두 만족하는 $CandidatePr$ 이다. $CandidatePr$ 은 후보 seam의 원소가 될 수 있는 픽셀의 집합이므로 $CandidatePr$ 내에서 후보 seam의 픽셀이 결정된다. 여기서 픽셀을 결정할 때에는 이웃하는 프레임 간의 상관성을 이용하여 $E(C_{ref})$ 의 에너지 값과 가장 유사한 픽셀을 p_{n+1} 로 결정한다. 에너지 함수로는 spatial forward energy(SFE)[6]와 제안하는 temporal forward energy(TFE)를 사용하였다. SFE 는 공간영역에 대한 에너지를 구하는 함수로 세로방향의 심을 구할 때에 식 (9)와 같이 표현된다.

$$\begin{aligned}
 SFE_{left-up}(i, j) &= |p(i+1, j) - p(i-1, j)| \\
 &\quad + |p(i, j-1) - p(i-1, j)|, \\
 SFE_{up}(i, j) &= |p(i+1, j) - p(i-1, j)|, \\
 SFE_{right-up}(i, j) &= |p(i+1, j) - p(i-1, j)| \\
 &\quad + |p(i, j-1) - p(i+1, j)|.
 \end{aligned} \quad (9)$$

여기서 $p(i, j)$ 는 (i, j) 위치의 픽셀을 의미한다. TFE 는 시공간영역에 대한 에너지를 구하는 함수로 식 (9)와 동일하게 세로방향의 심에 대해 값을 구할 경우에 식 (10)과 같이 표현된다.

$$\begin{aligned}
 TFE_{left-down}(i,j) &= |p(i-1,j,n-1) - p(i,j,n)| \\
 &\quad + |p(i+1,j,n-1) - p(i+1,j,n)|, \\
 TFE_{down}(i,j) &= |p(i-1,j,n-1) - p(i-1,j,n)| \\
 &\quad + |p(i+1,j,n-1) - p(i+1,j,n)|, \\
 TFE_{right-down}(i,j) &= |p(i-1,j,n-1) - p(i-1,j,n)| \\
 &\quad + |p(i+1,j,n-1) - p(i,j,n)|. \tag{10}
 \end{aligned}$$

여기서 $p(i,j,n)$ 은 n 번째 프레임의 (i,j) 위치에 존재하는 픽셀을 의미한다. SFE 와 TFE 를 이용하여 seam을 구성하는 p_{n+1} 이 도출되는 과정은 식 (11)과 같이 표현된다.

$$p_{n+1} = \arg \left[\min_{\alpha \in CandidatePx} (|SFE(\alpha) + TFE(\alpha)|) \right]. \tag{11}$$

p_{n+1} 은 공간연결성과 시간연결성을 동시에 고려한 에너지 함수가 최소값을 가질 때의 픽셀위치로 결정되며, seam은 식 (11)을 p_{height} (세로 방향 seam)까지 적용하여 생성된다. 최종 결과 영상은 앞선 과정으로 필요한 개수만큼 구해진 seam들을 이용하여 생성된다. 영상의 크기를 줄일 때는 seam의 순서대로 해당 seam을 구성하는 픽셀을 삭제하고, 반대로 영상의 크기를 늘릴 경우에는 seam의 순서대로 seam을 구성하는 픽셀의 위치에 픽셀을 추가한다.

IV. 실험

본 논문에서 제안하는 시스템의 성능평가는 bicubic, 기존의 Seam Carving(Conventional Seam Carving, CSC)^[6], 제안한 기법을 컬러영상에 적용하여 비교하는 실험으로 평가하였다. 실험은 영상처리에 널리 알려진 352×288

의 해상도를 가지는 CIF(common intermediate format) 컬러 영상 300프레임에 대해 수행하였으며, 가로방향으로 약 30%를 확대하여 458×288의 해상도를 가지는 결과영상을 얻었다. 또한 각각의 알고리즘이 수행되는데 필요한 실험 시간과 프레임당 메모리 사용량의 평균을 측정하였다. 실험과정에 발생될 실험의 오차를 줄여 신뢰도를 높이기 위해 동일한 실험 과정을 반복하였고, 반복과정을 통해 얻은 다수의 데이터를 평균하여 사용하는 방법을 적용하였다.

그림 8은 “Coastguard”(79th frame), “Mother & Daughter”(60th frame), “Soccer”(185th frame)에 대해 원본영상과 각각의 알고리즘이 적용된 결과영상을 비교하였다. 그림 8에서 나타나듯 Bicubic 기법은 영상의 모든 부분에 대해 동일한 크기로 픽셀을 보간하여 사용하기 때문에 모든 영상에 대해 전체적으로 주요 콘텐츠가 늘어나는 왜곡 현상이 발생한다. 이는 콘텐츠를 고려하지 않고 영상을 확대해서 생기는 문제로 모든 보간법이 이와 같은 문제를 내재하고 있다. CSC은 주요 콘텐츠를 고려하여 seam을 구하고 추가하여 콘텐츠가 보존되는 방법이지만, 동영상에서 적용될 경우에 프레임간의 연속성이 없고 seam이 주요 콘텐츠를 지나가는 현상이 발생되어 왜곡이 발생할 수 있다. 반면에 제안하는 고속 리타겟팅 기법은 seam을 구할 때에 프레임간의 시간적인 연결성도 고려되어 영상에 떨림현상이 없고, 주요 콘텐츠가 비교적 잘 보존되어 영상의 왜곡이 적게 나타나 원본 영상에 가까운 화질의 성능을 보여준다.

표 1은 제안하는 기법과 기존 알고리즘들의 실행시간(runtime)과 프레임당 평균 메모리 사용량(average for memory consumption of frame)을 측정한 성능비교를 나타낸다. 표 1의 결과에서 나타나듯 제안하는 기법은 CSC과 비교하여 약 5배 빠른 고속연산이 가능하고, 약 3배 적은 메

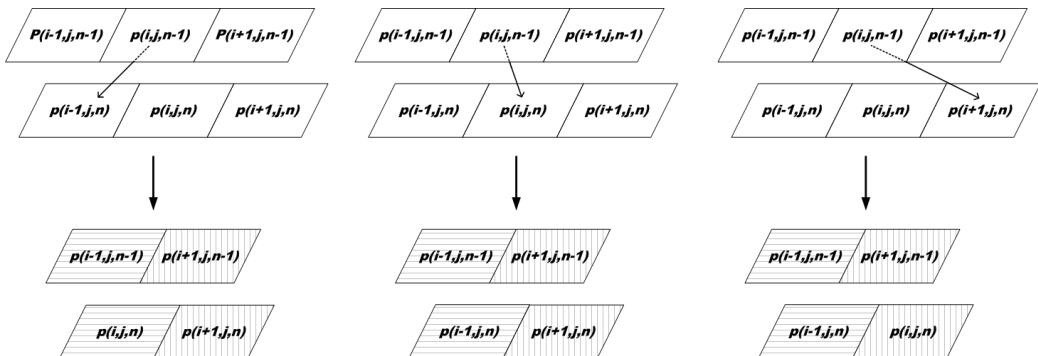


그림 7. $n-1$ 프레임과 n 프레임간의 temporal forward energy
 Fig. 7. Temporal forward energy between $(n-1)$ th frame and (n) th frame

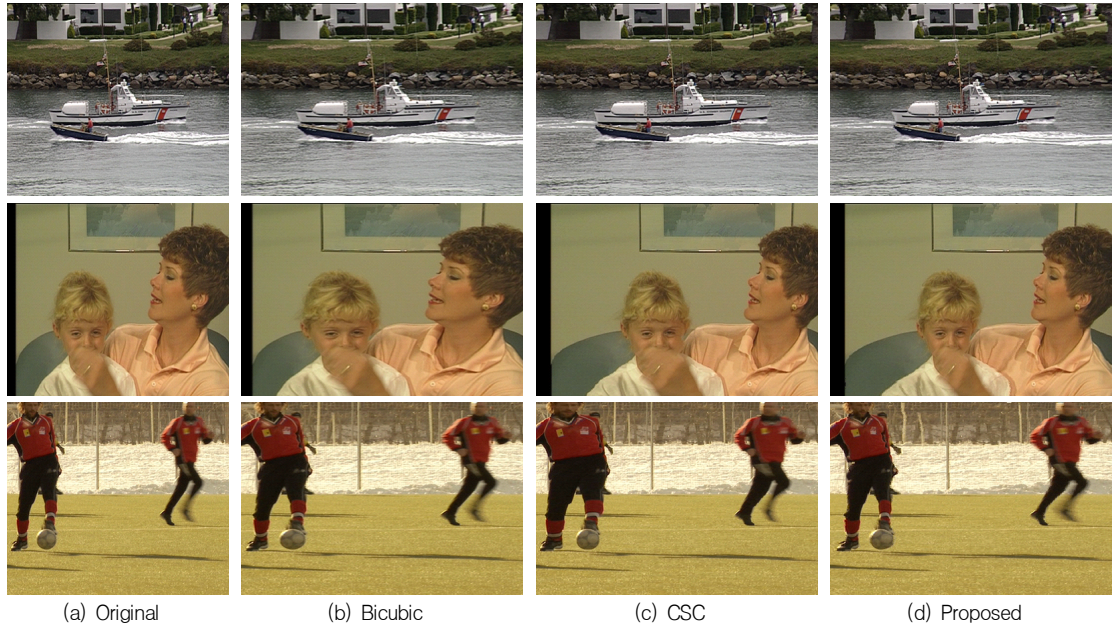


그림 8. 동영상 리타겟팅 기법들의 주관적 화질 비교
 Fig. 8. Subjective image quality comparison of video retargeting methods

모리의 사용만으로 고속 동영상 리타겟팅이 가능하다. 또한 16개의 픽셀을 참조하여 곱셈연산으로 계산되는 고차원 보간법(high-order interpolation)인 Bicubic보다는 연산량이 많아 실행시간이 다소 길지만, AMCF는 이전 프레임의 seam 정보만을 참조하여 연속하는 프레임의 seam을 구하기 때문에 근접한 수치로 나타난다.

나타나듯 CSC보다 제안하는 기법의 주파수 변화의 형태가 원본영상의 주파수변화와 유사함을 알 수 있다. 이는 제안하는 기법의 결과영상이 원본영상과 근접한 프레임의 주파수변화를 가진다는 것을 의미하고, 원본영상에 나타난 주요 컨텐츠가 결과영상에서도 잘 지켜지고 있음을 의미한다.

표 1. 동영상 리타겟팅 기법들의 성능 비교
 Table 1. Performance comparison of video retargeting methods

	Runtime(s)	AMCF(KB)
Bicubic	3.51	526
CSC	72.80	2232
Proposed	15.23	822

그림 9는 실험에서 비교대상으로 사용된 보간법인 bicubic을 제외하고 CSC와 제안하는 기법의 “Coastguard” 결과 영상을 차영상의 비교로 보여준다. 여기서 차영상은 역변환 연산(inverse operation)을 통해 재표현된 결과이다. 동영상에서 인접한 프레임간의 차영상은 시간축으로 존재하는 동일한 위치의 픽셀에 대한 주파수변화량으로 정의할 수 있다. 그림 9에

V. 결론

본 논문에서는 주요 컨텐츠를 보존하는 고속 동영상 리타겟팅 기법에 대해 제안하였다. 제안하는 방법은 scene이 시작되는 프레임에 대해서는 적응적 임계값과 공간의존도를 이용하여 계산된 누적 에너지로부터 복수개의 seam을 추출하여, 누적 에너지의 연산 횟수를 줄임으로써 처리 속도를 향상시킨다. 연속하는 프레임에 대해서는 이전 프레임 seam 정보를 참조하여 seam을 계산함으로써 처리 속도를 향상시키고, 결과 영상은 scene 전체 영상을 분석하지 않고도 컨텐츠의 떨림 현상이 존재하지 않게 도출된다. 제안한 고속 컨텐츠 인식 동영상 리타겟팅 기법은 영상의 주요 컨텐츠를 최대한 보존하면서도 기존의 Seam Carving 기법에 비하여 빠른 처리 속도를 가졌으며, 메모리 사용량도 현저히 적게 나타난다. 따라서 제안하는 시스템은 메모리 자원에 제약이 있고 실시간처리가 가능해

야만 하는 처리 환경인 모바일 단말과 같은 시스템에서 좋은 화질의 리타겟팅 영상을 얻기 위해 사용될 수 있다.

참고문헌

[1] K. T. Gribbon and D. G. Bailey, "A Novel Approach to Real-time Bilinear Interpolation", Proc. Second IEEE International Workshop on Electronic Design, Test and Application, Perth, Australia, pp. 126, 2004.

[2] R. Keys, "Cubic convolution interpolation for digital image processing", IEEE Transactions on Acoustics Speech And Signal Processing, Vol. ASSP-29, No. 6, pp. 1153-1160, 1981.

[3] I. S. Amrutha and S. S. Shylaja and S. Natarajan and K. N. Balasubramanya Murthy, "A smart automatic thumbnail cropping based on attention driven regions of interest extraction", Proc. ICIS on Computational Attention and Applications, 2009.

[4] M. Nishiyama and T. Okabe and Y. Sato and I. Sato, "Sensation-based photo cropping", Proc. ACM on Multimedia, pp. 957-962, 2009.

[5] F. Liu and M. Gleicher, "Automatic Image retargeting with fisheye-view warping," ACM Multimedia, pp. 153-164, Oct. 2005.

[6] S. Avidan and A. Shamir, "Seam Carving for Content-Aware Image Resizing," ACM Trans. on Graphics, Vol. 26, Issue 3, July 2007.

[7] M. Rubinstein and A. Shamir and S. Avidan, "Improved seam carving for video retargeting," ACM Transactions on Graphics, Vol. 27, Issue 3, Aug. 2008.

[8] B. Chen and P. Sem, "Video Carving," Proc. Eurographics, Short Papers, Apr. 2008.

[9] N. Otsu, "A threshold selection method from gray-level histogram", IEEE Transactions on System Man Cybernetics, Vol. SMC-9, No. 1, pp. 62-66, 1979.

[10] Z. Liu, H. Yan, K. N. Ngan, Z. Zhang and L. Shen, "Adaptive image retargeting using saliency-based continuous seam carving,"

Optical Engineering, Vol. 49, Issue 1, pp. 017002-017002-10, Jan. 2010.

[11] Y. Wang, J. Hsiao, O. Sorkine and T. Lee, "Scalable and Coherent Video Resizing with Per-Frame Optimization," ACM Transactions on Graphics, Vol. 30, No. 4, Article 88, July 2011.

[12] K. B. Kim and H.J. Lee, "Image Magnification using Fuzzy Method for Ultrasound Image of Abdominal Muscles", Journal of the Korea Society of Computer and Information, Vol. 16, No. 4, pp. 23-28, Apr. 2011.

[13] S. M. Jung, "Quality improvement scheme of magnified image by using gradient information between adjacent pixel values", Journal of the Korea Society of Computer and Information, Vol. 17, No. 2, pp. 59-67, Feb. 2012.

저자 소개



박대현

2007: 강원대학교
컴퓨터정보통신공학과 공학사
2009: 강원대학교
컴퓨터정보통신공학과 공학석사
2009~현 재: 강원대학교
컴퓨터정보통신공학과 박사과정
관심분야: 영상 신호 처리, 컴퓨터비전
Email : dhpark509@kangwon.ac.kr



김윤

1993: 고려대학교 전자공학과 공학사
1995: 고려대학교 전자공학과 공학석사
2003: 고려대학교 전자공학과 공학박사
1995~1999: LG-Philips
LCD 연구소 주임연구원
2002~2003: (주)달리텍 선임연구원
2004~현 재: 강원대학교
컴퓨터정보통신공학과 교수
관심분야: 영상 신호 처리
멀티미디어 통신 및 처리
Email : yooni@kangwon.ac.kr



그림 9. 기존 Seam Carving과 제안하는 기법의 차영상 비교
Fig. 9. Comparison of difference image for the Conventional Seam Carving and the Proposed method