

논문 2005-42SP-5-8

보간 검색을 이용한 Panning 비디오 영상에서의 모자이크 기법

(Mosaic Technique on Panning Video Images using Interpolation Search)

장 성 갑*, 김 재 신**

(Sung-gab Jang and Jae-shin Kim)

요 약

본 논문에서 삼각대 위에 고정시킨 카메라를 회전시켜 취득한 비디오 시퀀스로부터 파노라마 영상을 구성하는 새로운 방법을 제안한다. 제안한 방법은 비디오 시퀀스 중에서 실제로 사용할 프레임은 자동으로 선택하는 단계와 파노라마 영상을 구성하는 모자이크 단계로 나누어진다. 파노라마 영상을 구성하는데 필요한 프레임은 연속된 프레임 간의 중첩 비율 정보를 이용한 보간 검색을 통해 선택된다. 이 방법은 카메라의 속도에 상관없이 적절한 프레임을 고속으로 검색할 수 있다. 모자이크 단계에서는 4쌍의 특징점을 사용하는 사영 변환식(projective transform)을 이용한다. 기존의 연구에서는 특징점을 화소 기반의 정보만을 이용하여 추출하였지만, 본 논문에서 제안한 방법은 특징점의 위치까지 고려하여 사영 변환식의 정확도를 높였다. 제안한 방법은 다양한 영상 시퀀스에 적용하여 실험하였고, 그 결과 모든 영상에 대해서 좋은 화질의 파노라마 영상을 생성하였다.

Abstract

This paper proposes a new method to construct a panorama image from video sequences captured by the video camcorder revolving on the center axis of the tripod. The proposed method is consisted of two algorithms; frame selection and image mosaics. In order to select frames to construct the panorama image, we employ the interpolation search using the information in overlapped areas. This method can search suitable frames quickly. We construct an image mosaic using the projective transform induced from four pairs of quasi-features. The conventional methods select feature points by using only texture information, but the presented method in this paper uses the position of each feature point as well. We make an experiment on the proposed method with real video sequences. The results show that the proposed method is better than the conventional one in terms of image quality.

Keywords : panorama image, interpolation search, video mosaics, projective transfer

I. 서 론

최근 디지털 카메라와 각종 영상 저작 도구가 널리 보급됨에 따라 일반인들도 모자이크를 이용한 파노라마 영상 구성에 많은 관심을 보이고 있다. 현재 일반인들이 사용하는 파노라마 영상 저작 도구는 정지 영상을 기반으로 하기 때문에 사용자가 영상간의 적절한 중첩

영역을 고려하여 사진 촬영을 해야 하는 번거로움이 있다. 이러한 번거로움은 디지털 캠코더를 이용하여 촬영하면 많은 양의 영상 정보를 얻을 수 있기 때문에 쉽게 해소할 수 있다. 그러나 '많은 프레임 중 어떤 프레임을 선택하느냐' 하는 또 다른 불편을 초래한다.

본 논문에서는 디지털 캠코더로 취득한 동영상상을 기반으로 파노라마 영상을 자동으로 구성하는 방법을 제안한다. 제안한 방법은 취득한 동영상으로부터 적절한 프레임을 선택하는 단계와 선택된 프레임들로부터 파노라마 영상을 생성하는 단계로 구성된다. 디지털 캠코더

* 정회원, ** 학생회원, 중앙대학교 첨단영상대학원
(Chung Ang University)

접수일자: 2005년2월2일, 수정완료일: 2005년6월1일

로 취득한 영상은 많은 양의 영상 정보를 얻을 수 있지만, 파노라마 영상을 구성하기 위해서는 취득한 프레임 중 일부만으로도 충분하기 때문에 취득한 모든 영상을 이용해서 파노라마 영상을 구성하는 것은 대단히 비효율적이다. 기존 방법에서는 취득한 영상에 대한 정확한 평가 없이 3-5 프레임의 간격을 두고 선택해서 사용하거나 모든 프레임에 대한 평가 함수를 계산하여 프레임을 선택했다.^{[2][3]} 이러한 방법들은 카메라의 정확한 회전 속도에 대한 고려가 없었던 점에서, 그리고 모든 프레임을 일일이 검토한다는 점에서 비효율적이다. 본 논문에서는 파노라마 영상을 구성하는 효율적인 프레임 선택 알고리즘을 제안한다. 이 알고리즘은 영상간의 전역 정합 (global registration)을 이용하여 중첩 영역을 계산하고 보간 탐색을 통하여 카메라 회전 속도를 예측하여 프레임을 선택한다.

보간 검색을 통해서 적절한 프레임을 선택한 뒤 사영 변환식을 이용해서 파노라마 영상을 구성한다. 사영 변환식은 논문 [1]에서 성능을 입증한 유사 특징점을 좀더 개선한 특징점을 이용하여 계산한다. 기존 연구에서는 단 4쌍의 특징점만으로 영상간의 사영 변환식을 계산했기 때문에 빠르고 정확한 모자이크 결과를 보여주었지만, 4쌍의 특징점이 중앙에 집중되어 추출되면 사영 변환식의 정확성이 떨어지는 단점이 있다. 따라서 본 논문에서는 특징점 간의 거리를 고려한 가중치 함수가 포함된 새로운 특징점을 정의한다. 이러한 특징점은 영상 내에서 모서리에 가까운 쪽에서 추출되기 때문에 사영 변환식의 오차가 감소한다.

본 논문의 구성은 서론에 이어 II장에서 시스템 설계 및 프레임 선택 알고리즘을 설명한다. 그리고 III장에서 개선된 유사 특징점을 정의하는 방법을 설명하고, 유사 특징점을 이용해서 모자이크 영상을 구성하는 과정을 설명한다. 그리고 IV장에서 실험 결과 및 분석을 기술한 뒤 V장에서 결론을 맺는다.

II. 시스템의 설계 및 최적 프레임 선택

본 시스템의 목적은 디지털 캠코더로 촬영한 동영상으로부터 파노라마 영상을 구성하는 것이다. 그림 1에서 보는 바와 같이 파노라마 영상은 카메라를 회전시켜 취득한 영상들 간의 사영 관계를 해석하여 영상들을 합성한 것이다. 그림 2는 파노라마 영상을 자동으로 생성하기 위한 시스템 블록도를 나타내고 있다. 디지털 캠코더로 촬영한 영상은 IEEE 1394 보드를 이용하여 취

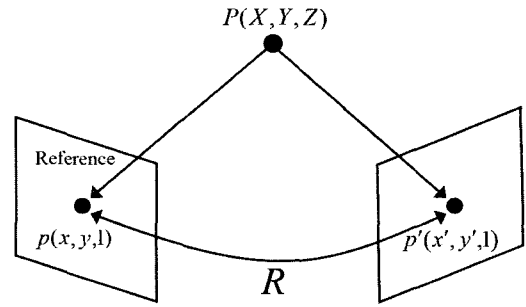


그림 1. 영상간의 투영관계
Fig. 1. Projective relation of images.

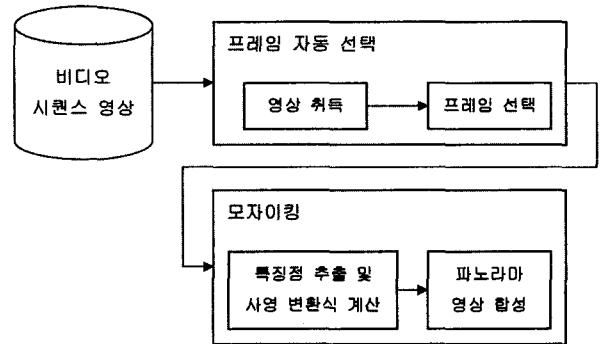


그림 2. 시스템 구조
Fig. 2. System overview.

득한다. 각각의 프레임은 비트맵 형식으로 저장되며, 연속된 프레임간의 중첩 영역 비율 정보에 대한 보간 검색을 통해 적절한 프레임을 선택한다. 이렇게 선택된 프레임에 대해서 특징점을 추출하고, 그 대응점을 검출하여 사영 변환식을 계산한다. 그리고 계산된 사영 변환식을 이용해서 최종 파노라마 영상이 생성된다.

디지털 캠코더의 경우 초당 최소 15 프레임이상의 영상을 촬영한다. 따라서 파노라마 영상을 구성하기 위해 디지털 캠코더를 삼각대에 고정된 뒤, 360도 횡 방향 촬영을 하면 수백 장의 영상이 취득된다. 수백 장의 영상을 모두 사용하여 모자이크를 하면 엄청난 양의 계산이 필요하므로 매우 비효율적이다. 따라서 취득한 수백 장의 영상 중에서 실제로 사용할 프레임을 선택하는 작업이 필요하다. 이 작업은 카메라의 회전 속도를 고려해야 하는데 사용자의 입력 없이 영상 정보만으로 회전 속도를 알기 위해 전역 정합을 이용한다. 전역 정합을 이용하면 영상간의 중첩 영역을 추출할 수 있기 때문에 프레임 간의 인덱스 (index) 차이와 중첩 영역 비율 값을 이용해 카메라 회전 속도를 예측할 수 있다.

1. 전역 정합

연속된 두 영상의 중첩 영역은 Kuglin 과 Hines 가



그림 3. 전역 레지스트레이션의 결과로부터 추출된 중첩 영역

Fig. 3. Result of overlap area by global registration.

제한한 주파수 공간에서 영상의 위상 상관 관계를 이용해서 추출할 수 있다. 이 방법은 2차원 푸리에 변환을 취한 두 영상들 사이의 변위(translation)를 계산하여 중첩 영역을 검출한다. 이 방법은 계산량이 적고, 잡음에 강건한 성질을 갖는다.^[8]

두 영상을 각각 $f_1(x, y)$, $f_2(x, y)$ 그리고 각각의 푸리에 변환식을 $F_1(u, v)$, $F_2(u, v)$ 라 하자. 이 두 영상이 식 (1) 과 같은 관계를 갖는다면, F_1 , F_2 사이의 관계는 식 (2) 와 같다.

$$f_2(x, y) = f_1(x + x_0, y + y_0) \quad (1)$$

$$F_2(u, v) = F_1(u, v) \cdot e^{-j2\pi(ux_0 + vy_0)} \quad (2)$$

식(2)로부터 두 영상의 정규화된 cross-power spectrum 을 정의 할 수 있다.

$$\frac{F_1(u, v)F_2(u, v)^*}{|F_1(u, v)F_2(u, v)|} = e^{-j2\pi(ux_0 + vy_0)} \quad (3)$$

식(3)을 역 푸리에 변환(inverse Fourier transform) 하면 공간상에서 임펄스 함수(impulse function)를 나타낸다. 이 임펄스의 위치가 두 영상의 변위(displacement) 정보를 나타낸다. 이 변위 정보를 기준으로 중첩 비율을 정의하였다. 그림 3의 밝은 부분은 전역 정합의 결과로 추출된 중첩 영역이다. 그림 3의 왼쪽 영상을 기준 영상(reference image), 오른쪽 영상을 대상 영상(target image) 으로 지칭하기로 한다.

2. 보간 검색을 이용한 최적 프레임 선택

전역 정합은 중첩 비율이 최소 50% 이상일 때 비교적 정확하게 추출된다.^[6] 본 논문에서는 보다 강건한 모자이킹을 위해 70% 의 중첩 비율을 가지는 프레임을 이용한다. 따라서 프레임 선택의 목표는 디지털 캠코더

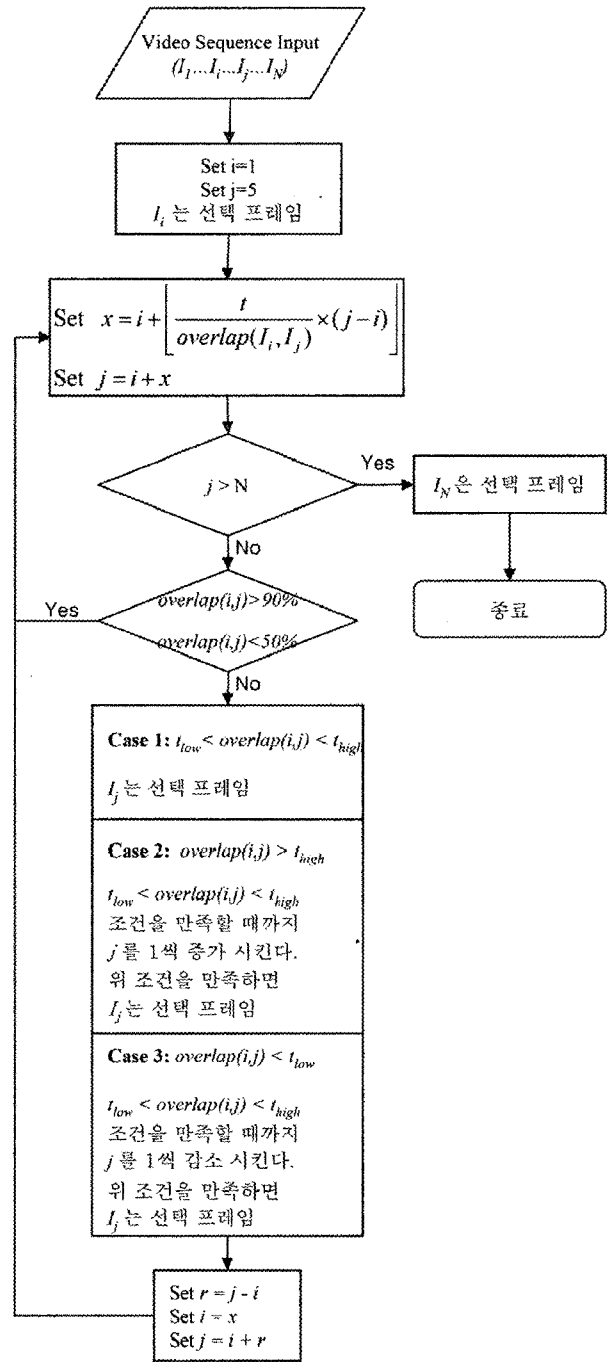


그림 4. 프레임 선택을 위한 보간 검색 알고리즘

Fig. 4. Interpolation search algorithm for frame selection.

를 통해 취득한 수백 장의 영상 중에 첫 번째 프레임울 기준으로 매 70% 중첩 영역을 가지는 프레임을 선택하는 것이다. 이 작업을 위해 순차적으로 모든 프레임간의 중첩 영역을 계산하는 것은 많은 시간을 요구한다. 따라서 보간 검색을 이용하여 카메라 회전 속도를 예측하여 보다 빠른 시간 안에 프레임을 선택한다. 보간 검색은 데이터가 순차적으로 정렬되어 있을 때 임의의 두

	i							x	
Index	1	2	3	...	25	26	27	28	29
중첩비율(%)	99	94	87		78	75	70	65	54

그림 5. 프레임 선택 알고리즘
Fig 5. Frame selection algorithm.

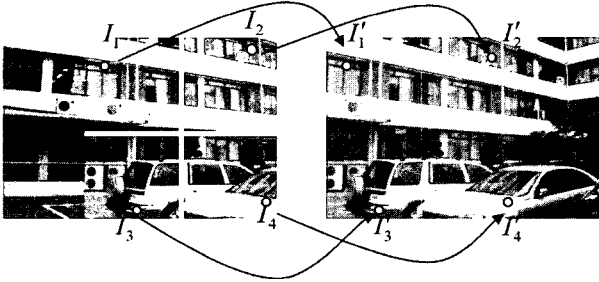


그림 6. 유사 특징점 및 대응점 검출
Fig. 6. Extraction of quasi-feature points and correspondence points.

개의 데이터를 이용하여 찾고자 하는 데이터의 인덱스를 예상하는 검색 방법이다. 여기서 사용하는 보간 검색은 흔히 사용하는 임계값 (threshold)이 하나인 보간 검색이 아닌 임계값이 두 개 (high threshold, low threshold)인 응용된 보간 검색이다. 임계값이 두개인 보간 검색은 보다 빠른 프레임 선택 시간을 보증하는 반면, 70%에 가장 가까운 중첩 영역을 가지는 프레임을 선택하지 못한다. 그러나 모자이크 영상을 구성함에 있어서 중첩 영역 비율이 50% 이상이면 영상을 구성하는데 큰 무리가 없기 때문에 큰 문제가 되지 않는다. 또한 두 임계값을 70%에 가깝게 설정함으로써 모자이크 영상 구성의 강건함과 효율성을 잃지 않도록 하였다.

그림 6 은 적절한 프레임을 선택하는 보간 검색 알고리즘이다. t_{low} 는 low threshold 값인 65 이고 t_{high} 는 high threshold 값인 75 이다. t 는 threshold 값인 70 이다. $overlap(I_i, I_j)$ 는 두 영상간의 중첩 비율을 계산하는 함수이다. 카메라 회전 속도에 상관없이 적절한 프레임을 찾기 위해 식(4) 와 같은 보간 검색식을 사용하였다. i 는 최근 선택 프레임 인덱스이고 j 는 $i+r$ 이다. r 은 알고리즘 수행 과정에서 계산된다. 다음 식은 프레임 인덱스 차에 따른 중첩 비율을 계산하여 적절한 중첩 비율을 가지는 프레임까지의 간격인 x 를 추정한다. 현재 프레임 I_i 와 I_{i+x} 와의 중첩 영역 비율을 계산하고 3가지 경우에 따른 처리를 하여 최적의 프레임을 선택한다. 만약 중첩 영역 비율이 임계값과 심하게 차이가 나면 식(4) 를 이용하여 x 를 다시 계산한다.

$$x = i + \left\lfloor \frac{t}{overlap(I_i, I_j)} \times (j - i) \right\rfloor \quad (4)$$

그림 5는 알고리즘에 대한 이해를 돕기 위한 그림이다. 그림에서 음영이 들어간 26, 27, 28 인덱스가 최적의 중첩비율을 갖는 프레임 인덱스이다. 현재 프레임이 i 인덱스일 때 식 (4) 를 이용하여 x 값을 구한 뒤 $overlap(i, i+x)$ 값이 65% 에서 75% 사이 값이면 그 프레임을 선택한다.

III. 유사 특징점 정의 및 대응점 검출

그림 1의 두 영상관의 관계는 $p = H \cdot p'$ 의 사영 변환식 H 로 표현할 수 있다. 사영 변환식을 계산하기 위해서는 먼저 4개의 특징점과 대응점을 기준 영상의 중첩 영역에서 추출해야 하며, 이 특징점 중 임의의 세 점은 동일 직선상에 존재하지 않아야 한다.^[7] 따라서 추출된 기준 영상의 중첩 영역을 4개의 부 영역으로 분할한 뒤, 각 영역에서 유사 특징점을 추출한다. 그리고 이에 대한 대응점을 대상 영상에서 찾아 4쌍의 특징점을 이용, 사영 변환식을 유도한다. 본 장에서는 개선된 유사 특징점 정의 방법과 대응점 검출에 대해 기술한다. 그림 6 은 대응관계를 표시한 그림이다. I_n 은 유사특징점이고 I'_n 대응점을 나타낸다.

1. 유사 특징점 정의

유사 특징점은 화소의 그레이 레벨(grey-level) 분포를 기반으로 하는 특징점이다. 즉, 충분한 텍스처 정보를 갖는 화소의 블록은 다른 블록에 비해 변화량이 크기 때문에 블록 내 분산값이 크다. 기존의 유사 특징점은 영상 내에서 분산의 크기가 최대인 블록의 중심 화소로 정의하였다. 이러한 점을 특징점으로 선택하면 보다 쉽게 대응점을 검출할 수가 있다. 그러나 이러한 특징점에 대한 대응점 검출 시, 어느 정도의 오차를 수반하기 때문에 정확한 사영 변환식을 유도하는 데에 한계가 있다. 이러한 오차는 실제 정확한 대응점과 추출된 대응점간의 거리로 표현되는데, 이 거리가 클수록 오차가 커져서 정확한 사영 변환식을 유도하지 못하게 된다. 이러한 오차는 특징점이 중앙에서 추출되었을 때 그 영향이 더 커지게 된다.^[7]

따라서 본 논문에서는 4개의 특징점만을 이용한 직관적이고 빠른 사영 변환식 계산법을 유지하면서, 오차

의 영향을 줄이기 위해 특징점의 위치까지 고려하는 새로운 특징점을 제안한다. 영상에서 최대 분산값을 갖는 특징점과 비슷한 분산 값을 갖는 화소는 여러 개 존재한다. 이러한 특징점도 사영 변환식에 충분히 사용할 수 있다. 따라서 여러 개의 특징점 중에서 중첩 영역의 모서리에 가까운 특징점을 최종 특징점으로 선택한다면 좀 더 정확한 사영 변환식을 얻을 수 있다. 특징점을 추출하기 위해 중첩 영역을 4등분 하여 4개의 부 영역을 정의하고 각 부 영역에서 15x15 블록을 이용하여 특징점을 추출하였다. 우리는 논문 [1]의 유사 특징점에 가중치 함수가 추가된 개선된 유사 특징점을 식(5)와 같이 정의한다. 식(5)는 i 번째 부 영역, k 번째 블록에서 선택된 유사 특징점 q_i 이다.

$$q_i = \arg \max[\sigma_{k,i}^2(1-w(x,y))], 0 \leq i \leq 3$$

$$\sigma_{k,i}^2 = \sum_k \left[\sum_{g=0}^{G_{MAX}} g \cdot h_g - M_k^2 \right]$$

$$w(x,y) = 0.1 \times (dis(x,y) / dis(width,height)) \quad (5)$$

$\sigma_{k,i}^2, M_k$ 는 i 번째 부 영역에서 k 번째 블록의 분산과 평균을 각각 나타내고 h_g 는 그레이 레벨 g 에서 히스토그램이다. G_{MAX} 는 그레이 레벨의 최대값(255)이다. $w(x,y)$ 는 제안한 가중치 함수다. 그림 7에서 표시한 것과 같이 $dis(w,y)$ 는 현재 블록의 중심 좌표에서 해당 부 영역의 모서리까지의 거리이고, $dis(width,height)$ 는 i 번째 부 영역의 모서리 좌표에서 중첩영역의 중심 좌표까지의 거리이다.

$dis(x,y)/dis(width,height)$ 값이 모서리에 가까울수록 $w(x,y)$ 값은 작아지므로 q_i 에 정의된 $1-w(x,y)$ 가중치 값은 모서리에 가까울수록 커진



그림 7. 특징점 위치에 따른 가중치 함수
Fig. 7. Weight function depend on position of quasi-feature points.

다. 그러므로 같은 분산값을 갖는 특징점이라도 모서리에 가까운 점이 특징점으로 정의된다.

2. 블록 정합 알고리즘을 이용한 대응점 검출

기준 영상의 유사 특징점 4개가 정의되면 대상 영상에서 대응점을 찾는다. 대응점을 찾는 알고리즘은 카메라 왜곡이 고려된 가중치 함수를 포함한 블록 정합 알고리즘(BMA: Block Matching Algorithm)을 사용한다.^[1] 이 알고리즘은 블록 단위(B)로 화소를 비교하여 오차가 가장 적은 점을 대응점으로 검출한다.

$$E(d_x, d_y) = \left[\sum_{(x,y) \in B} |s(x,y,k) - s(x+d_x, y+d_y, k+1)| \cdot w_{ij} \right] \quad (6)$$

$s(x,y,k)$ 는 k 프레임의 (x,y) 위치에서 화소값이고, d_x, d_y 는 블록의 중심에서 변위를 나타낸다. 최적의 변위는 최소 E 를 나타내는 (d_x, d_y) 이고, 식(7)처럼 표현된다.

$$[\hat{d}_x, \hat{d}_y] = \operatorname{argmin} E(d_x, d_y) \quad (7)$$

IV. 실험 결과

제안한 알고리즘은 IBM PC Pentium4 2.0GHz 에서 실험하였으며, 영상은 Sony VX2000 DV Camcorder로 취득하였다. 실험 영상은 프레임당 360x240 크기를 가진다.

1. 프레임 선택 결과

표 1과 표 2는 프레임 선택 알고리즘 결과를 분석한 것이다. 표 1에서 '정지 영상 기반'은 프레임 자동 선택 알고리즘 없이 선택된 프레임만으로 파노라마 영상을 구성한 시간이고 '비디오 영상 기반'은 본 논문에서 제안한 파노라마 영상 구성 시간이다. 정지 영상 기반의 방법에서는 비디오 영상 기반에서 선택된 프레임은 직접 선택하여 수행 시간을 측정하였다. 표 1의 결과에서 알 수 있듯이 제안한 프레임 선택 알고리즘을 이용할 경우 정지 영상에서 사용하는 영상 수의 수십 배의 영상을 대상으로 파노라마를 구성함에도 불구하고 정지 영상 기반의 방법보다 2배 정도의 시간만을 요구하는 것을 알 수 있다. 따라서 본 논문에서 제안한 알고리즘을 통해 사용자는 중첩 영역에 대한 고려 없이 비디오로 영상을 취득하여 빠른 시간 안에 파노라마 영상을

효율적으로 구성할 수 있다.

표 2 는 비디오 panning 속도에 따른 선택 프레임 번호와 선택된 프레임 간의 중첩 비율을 나타낸 것이다. 카메라 회전을 특별한 장치 없이 사용자가 직접 수행하

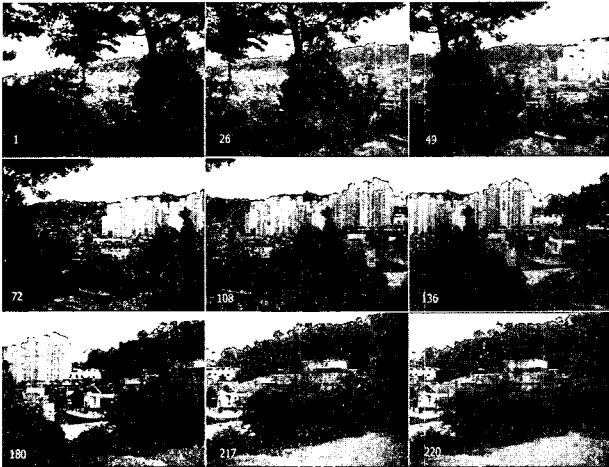


그림 8. 영상 1에 대한 프레임 선택 결과
Fig. 8. Result of frame selection for image 1.

표 1. 수행 시간 비교
Fig. 1. Comparison of execution time.

영상 1	
선택 프레임 수	9/220
걸린 시간	정지 영상 기반: 14초 비디오 영상 기반: 25초
영상 2	
선택 프레임 수	9/203
걸린 시간	정지 영상 기반: 14초 비디오 영상 기반: 28초
영상 3	
선택 프레임 수	8/121
걸린 시간	정지 영상 기반: 11초 비디오 영상 기반: 22초
영상 4	
선택 프레임 수	12/228
걸린 시간	정지 영상 기반: 20초 비디오 영상 기반: 41초

표 2. Panning 속도에 따른 선택 프레임 번호 및 선택 프레임 간의 중첩 비율
Fig. 2. Selected frame index and the ratio of the overlapped region between selected frames for the panning speed.

영상	Panning 속도	내용	선택된 프레임과 프레임 간 중첩 비율 비교 데이터											
			선택 프레임 번호	1	26	49	72	108	136	180	217	220		
영상 1	느림	선택 프레임 번호	1	26	49	72	108	136	180	217	220			
		중첩 비율(%)		69	69	73	67	66	70	65				
	중간	선택 프레임 번호	1	11	18	27	39	48	62	72	75			
		중첩 비율(%)		64	70	71	69	68	71	71				
	빠름	선택 프레임 번호	1	4	6	8	11	14	17	21	24	25		
		중첩 비율(%)		68	75	77	75	70	71	75	75			
영상 2	느림	선택 프레임 번호	1	44	71	89	109	129	149	169	197	203		
		중첩 비율(%)		73	76	71	72	70	71	70	69			
	중간	선택 프레임 번호	1	15	24	30	36	43	49	55	64	68		
		중첩 비율(%)		74	74	72	73	71	74	72	70			
	빠름	선택 프레임 번호	1	6	10	12	15	17	19	22	23			
		중첩 비율(%)		72	65	69	64	74	72	70				
영상 3	느림	선택 프레임 번호	1	35	52	70	84	107	121					
		중첩 비율(%)		65	70	71	70	69						
	중간	선택 프레임 번호	1	12	17	22	27	33	39	41				
		중첩 비율(%)		67	73	73	73	73	73					
	빠름	선택 프레임 번호	1	4	6	9	11	13	14					
		중첩 비율(%)		73	71	68	73	71						
영상 4	느림	선택 프레임 번호	1	39	65	88	113	144	153	171	189	206	228	
		중첩 비율(%)		69	73	69	70	69	72	67	70	70		
	중간	선택 프레임 번호	1	14	22	29	38	48	51	55	59	65	70	76
		중첩 비율(%)		68	74	70	69	70	76	76	74	75	71	
	빠름	선택 프레임 번호	1	5	8	10	13	16	18	21	23	25	26	
		중첩 비율(%)		71	72	75	70	72	70	68	72	73		

기 때문에 panning 속도를 상대적으로 표현하였다. Panning 속도가 느린 경우에는 선택 프레임 번호간의 간격이 넓고 panning 속도가 빠른 경우에는 선택 프레임 번호간의 간격이 좁은 것을 볼 수 있다. 즉, 사용자의 특별한 입력 없이도 제안한 알고리즘에 의해 최적의 프레임을 적절한 간격으로 자동 선택하고 있다. 표에 나타난 중첩 비율은 이전 선택 프레임과의 중첩 비율이다. 모든 프레임에서 적절한 중첩 영역을 가지는 최적의 프레임을 찾는 것을 알 수 있다. 또한 panning 속도에 상관없이 비디오 파노라마 구성 시간이 표 1에서의 시간과 똑같이 측정되었다. 이것은 총 수행 시간이 주어진 비디오 영상의 양과 상관없이 최종 구성할 파노라마 영상의 크기에만 상관관계를 가짐을 보여주는 내용이다. 그림 8 은 프레임 선택 결과의 한 예이다.

2. 개선된 유사 특징점 결과분석

그림 9 는 참고문헌 [1] 에서 제안한 유사 특징점 추출 방법과 본 논문에서 제안한 유사 특징점 추출 방법을 비교한 그림이다. 그림 10 은 추출된 유사 특징점과 대응점으로 모자이킹을 한 결과이다. 그림 9(a) 를 보면 논문[1]의 방법의 경우 대응점이 정확하게 검출되었음에도 불구하고 I_1 특징점이 너무 오른쪽에서 추출되어 영상의 왼쪽 위 부분의 모자이킹 결과가 좋지 않음을 알 수 있다.(그림 10a) 이에 비해 그림 10(b) 를 보면 제안한 방법으로 특징점을 추출했을 경우 특징점 P_1 이 모서리에 가깝게 추출되어 정확하게 모자이킹 되었음을 알 수 있다.

제안한 알고리즘의 성능을 다양한 영상에 대해 객관적으로 평가하기 위해서 식 (8) 과 같은 오차 함수를 이용한다. $I(x,y)$ 는 영상의 밝기 값이다. 표 3는 오차 함수를 통해 계산된 성능 비교표이다. 비교 대상인 기존 방법은 논문 [1]에서 제안한 특징점 정의이다. 표 1 에서 선택된 프레임을 대상으로 오차 함수를 계산하였다. 표 3과 그림 11 에서 알 수 있듯이 제안된 방법이 이전 방법에 비해 성능이 향상되었음을 알 수 있다.

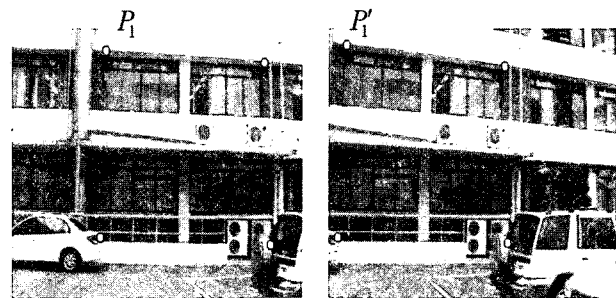
$$E = \frac{1}{N} \sum_R \{I(x,y) - I'(x,y)\}^2$$

$$R = \text{overlaparea} \tag{8}$$

그림 12 는 디지털 캠코더로 취득한 다양한 영상에 대해 제안한 알고리즘을 이용하여 비디오 파노라마 영상을 구성한 최종 결과를 보여주고 있다.

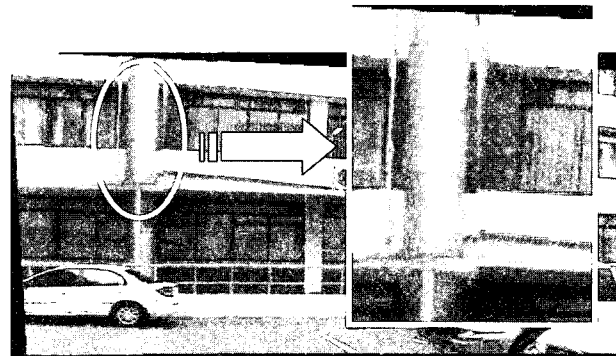


(a) 이전 방법

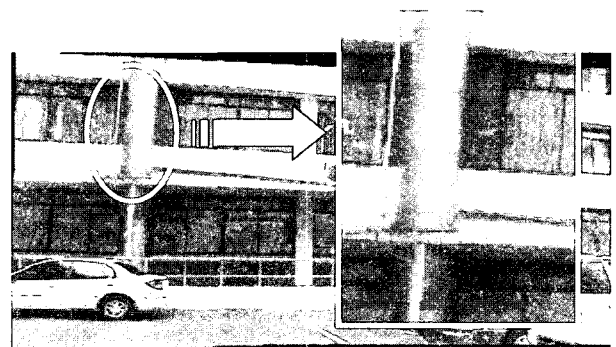


(b) 제안 방법

그림 9. 유사 특징점과 대응점 검출
Fig. 9. Extraction of quasi-feature points and correspondence points.



(a) 이전 방법



(b) 제안 방법

그림 10. 모자이킹 결과 비교
Fig. 10. Comparison of mosaicing results.

표 3. 기존 방법과 제안한 방법과의 오차의 비교 (단위: 식 (8) 의 MSE)
 Table 3. Comparison between previous and proposed method.

프레임 Index	영상 1		영상 2		영상 3		영상 4	
	기존 방법	제안한 방법	기존 방법	제안한 방법	기존 방법	제안한 방법	기존 방법	제안한 방법
1-2	1186.65	1159.68	1029.78	901.95	1577.24	1439.53	698.57	499.19
2-3	1186.12	1171.65	771.1	697.91	2498.57	2481.79	898.1	788.17
3-4	1197.35	1177.12	1068.52	1009.06	2175.9	2175.9	809.05	799.95
4-5	1352.35	1352.35	1414.76	1414.76	1483.75	1407.08	1149.29	1150.13
5-6	1754.08	1729.26	1865.08	1865.81	1054.38	985.42	1356.93	1356.93
6-7	1388.75	1386.06	1194.01	1200.06	991.94	971.81	1322.73	1304.2
7-8	1097.24	1109.88	1329.56	1313.8	1011.91	1003.71	1142.17	1021.99
8-9	769.23	735.11	1215.98	1210.08	.	.	949.48	916.41
9-10	761.61	715.06

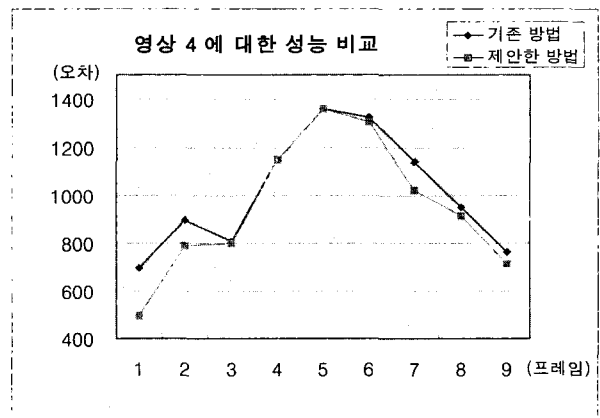
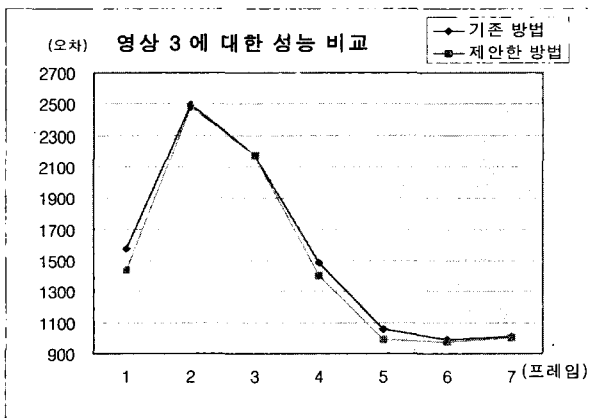
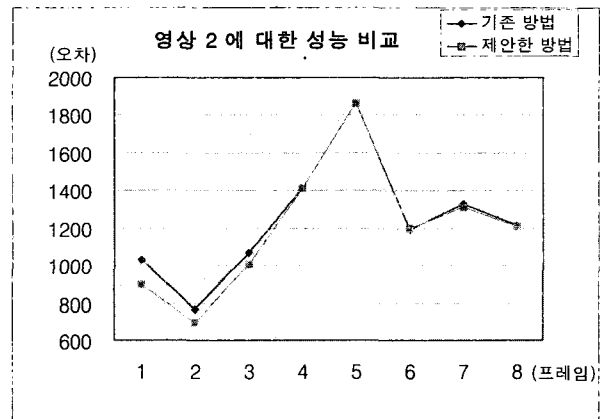
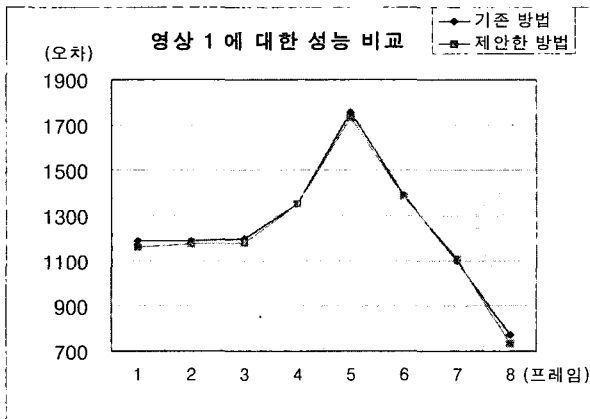


그림 11. 기존 방법과 제안한 방법과의 성능 비교 그래프
 Fig. 11. Performance comparison graph between previous and proposed method.



그림 12. 비디오 파노라마 구성 결과

Fig. 12. Results of construction video panorama.

V. 결 론

본 논문에서는 디지털 캠코더로 취득 된 수백 장의 영상 중에서 모자이킹에 사용할 일부의 영상을 취득하는 효율적인 방법과 모자이킹 결과 향상을 위한 유사 특징점 정의 방법을 새롭게 제안하였다. 일정 구역을 정해 보간 검색을 수행함으로써 프레임을 빠르게 검색할 수 있었다. 뿐만 아니라 검색 과정에서 적절한 중첩 영역을 가지는 프레임을 선택함으로써 이후 수행하는 모자이킹에서도 강건한 결과를 보여주었다. 또한 특징점의 위치를 고려한 가중치 함수를 적용한 유사 특징점 정의를 통해 특징점의 위치가 중앙으로 몰리지 않도록 함으로써 더욱 향상된 파노라마 영상을 얻었다. 본 알고리즘을 사용한 응용 프로그램 개발 시 일반인들 또는 파노라마 영상 기반 관광지 및 촬영지 검색 웹 사이트 개발자들이 정지 영상 기반 프로그램보다 쉽게 파노라마 영상을 구성할 수 있을 것으로 기대된다. 향후 파노라마 이미지를 볼 수 있는 네비게이션 시스템과 고정된 위치가 아닌 이동하면서 네비게이션 할 수 있는 시스템으로의 확장에 관한 연구가 진행 되어야 할 것이다.

참 고 문 헌

- [1] 김대현, 윤용인, 최중수, "유사특징점을 이용한 모자이킹 영상의 구성", 전자공학회 논문지, pp. 제38 권 SP편 7월호, 2001.
- [2] Xueying Qin, Eihachiro Nakamae, katsumi Tadamura, "Creating a Precise Panorama from Panned Video Sequence Images", Computer Graphics and Applications, 1999. Proceedings. Seventh Pacific Conference, pp 6-11, Oct. 1999.
- [3] J. S. Jimmy Li, Sharmil Randhawa, "Improved Video Mosaic Construction by Selecting a Suitable Subset of Video Images", Proceedings of the 27th conference on Australasian computer science - Volume 26, January. 2004.
- [4] R. Szeliski, "Video Mosaics for Virtual Environments", IEEE Computer Graphics and Applications, vol. 16, no. 2, pp 21-30, Mar. 1996.
- [5] L. G. Brown, "A Survey of Image Registration Techniques", ACM Computing Surveys, Vol.24, No.4, pp.325~376, Dec. 1992.
- [6] R. Szeliski, "Image Mosaicing for Tele-Reality applications", IEEE Workshop on Applications of Computer Vision (WACV'94), pp. 44-53, Dec. 1994.
- [7] R. Hartley and A. Zisserman "Multiple View Geometry in Computer Vision", Cambridge Univ. Press, Ch. 2,4
- [8] S. E. Chen, "QuickTime VR an image-based approach to virtual environment navigation", Computer Graphics (SIGGRAPH '95), pages 29-38, August 1995.

[1] 김대현, 윤용인, 최중수, "유사특징점을 이용한 모

[9] E. Trucco, A. Verri, "Introductory Techniques for 3D Computer Vision", pp. 123-138, Prentice Hall, Upeer Saddle River, 1998.

[10] Y. Huang, X. Zhuang, "An Adaptively Refined Block Matching Algorithm for Motion Compensated Video Coding", IEEE Trans. on Circuits/Systems for Video Tech. Vol.5, No.1 pp.56~59, Feb. 1995.

저 자 소 개



장 성 갑(정희원)
 1988년 중앙대학교 전자공학과
 학사 졸업
 1990년 중앙대학교 정보제어학과
 석사 졸업
 2000년 중앙대학교 영상공학과
 박사 졸업.

2000년~현재 중앙대학교 첨단영상대학원
 영상예술학과 조교수

<주관심분야 : 영상 처리, 영상 특수 효과>



김 재 신(학생회원)
 2004년 중앙대학교 컴퓨터 공학과
 학사 졸업.
 2004년~현재 중앙대학교 첨단
 영상대학원 영상공학과
 석사 과정.

<주관심분야 : 파노라마 영상, 스테레오 및 다 시
 점 영상 보간, 3D TV>