보로노이 거리(Voronoi Distance)정합을 이용한 영상 모자익

이칠우[†] · 정민영[††] · 배기태[†††] · 이동휘[††††]

요 약

본 논문에서는 손에 든 비디오 카메라로 촬영한 영상을 결합하여 대형 고해상도 영상을 생성하는 모자익 기법에 관해 기술한다. 기존의 특징점 기반 대응점 검색 기법들이 사람의 개입에 의해 영상을 정합하거나, 형태정보만 이용하는 방법을 사용 하는 것과는 달리 형태정보와 칼라정보를 모두 사용하여 사람의 개입이 없이 자동으로 정확한 중첩영역을 검색하는 방법을 제안한다. 제안한 방법은 인접하는 영상 간에 빠른 속도로 대응점을 추정하여 영상간의 초기 변환관계를 계산하는 보로노이 거리(Voronoi Distance)정합법을 이용하여 비슷한 형태를 가진 후보 영역들을 추출한 다음, 칼라 정보를 이용하여 최종 중첩영역을 찾는다. 이것은 영상내의 특징점을 기준으로 특징점 사이의 거리가 동일하도록 기준영상의 보로노이 평면(Voronoi Surface)을 생성하고 입력영상과 기준 영상간의 정합 부분에서 누적된 보로노이 거리를 최소화하는 대응점을 이항검색 기법으로 추출하는 방법이다. 추출된 계산된 초기 변환행렬은 Levenberg-Marquadt 방법을 통해 최적 변환행렬로 수정되고 이 변환행렬에 의해 영상이 합성되어진다.

Image Mosaicing using Voronoi Distance Matching

Chil-Woo Lee[†], Chong Min-Yenog[††], Ki-Tae Bae[†††] and Dong-Hwi Lee[††††]

ABSTRACT

In this paper, we describe image mosaicing techniques for constructing a large high-resolution image with images taken by a video camera in hand. we propose the method which is automatically retrieving the exact matching area using color information and shape information. The proposed method extracts first candidate areas which have similar form using a Voronoi Distance Matching Method which is rapidly estimating the correspondent points between adjacent images, and calculating initial transformations of them and finds the final matching area using color information. It is a method that creates Voronoi Surface which set the distance value among feature points and other points on the basis of each feature point of a image, and extracts the correspondent points which minimize Voronoi Distance in matching area between an input image and a basic image using the binary search method. Using the Levenberg-Marquadt method we turn an initial transformation matrix to an optimal transformation matrix, and using this matrix combine a basic image with a input image.

Key words: Image Mosaic, Vornnoi Surface, 이항 검색(Binary Search), Levenberg-Marquadt method, 영상합성, 영상기반 랜더링

본 연구는 한국과학재단 지정 전남대학교 "고품질 전기 전자 부품 및 시스템 연구센터"의 연구비 지원에 의해 수행되었음.

접수일 : 2003년 6월 24일, 완료일 : 2003년 8월 20일

[†] 종신회원, 전남대학교 컴퓨터공학과 부교수

[††] 정회원, 광주 여자 대학교 멀티미디어학과 조교수

[†††] 준회원, 전남대학교 컴퓨터공학과 대학원 박사수료

[††††] (주)버추얼미디어 부설기술연구소 연구원

1. 서 론

최근 영상 기반 랜더링 기술의 하나로 모자익 기술이 컴퓨터 비젼 분야에서 매우 흥미로운 테마로 자리잡아가고 있다. 이 기술은 임의의 시점에서 촬영한 영상들을 하나의 대형영상으로 통합시킴은 물론 계산에 의해 시점을 달리하여 영상을 표현할 수 있기 때문에, 건축물 내부와 같은 복잡한 장면들을 그래픽 모델을 사용하지 않고 간단하게 생성할 수 있다. 때문에 영상 생성이 필요한 각종 응용 시스템에서 많이 이용될 수 있다. 이 기술이 연구 초기에는 우주 및 항공 사진의 합성과 같이 제한적인 분야에서만 활용되었으나 실영상을 입체적으로 조작할 수 있게 되면서 컴퓨터 그래픽스의 페인팅 시스템, 가상현실 시스템, 영화제작 등 많은 분야에서 활용되고 있다[1].

모자익 알고리즘에 있어서 중요한 요소기술은 두 영상을 접합할 중첩영역을 찾는 기술과 중첩영역을 기준으로 여러 영상을 불연속점 또는 왜곡이 없이 매끄럽게 합성하는 기술이다. 복수의 영상간의 대응관계를 찾는 것은 스테레오 기하에 의한 3차원 정보 추출을 위해 많이 연구되어 왔으나 아직까지 정확히 일반화된 해법이 제시되어 있지 않다.

모자익을 위해서는 영상간의 상관을 이용하거나[2], 특징점간의 거리 제약[4] 또는 광류[8]를 이용하는 방법들이 있다. 영상간의 상관을 이용하는 방법은 중첩영역간의 단순 휘도차를 이용하기 때문에 직관적이고 구현이 쉬우나 영상의 조명 환경, 즉 영상전체의 밝기가 변하면 정확한 대응을 찾기가 어렵게 된다. 특징점 간의 거리 제약을 이용하는 방법[9]은 찾아낸 특징 점들이 영상 간에 서로 임의의 대응을 갖는다고 가정하고 그중 거리 오차가 가장 적은 결합을 영상 대응으로 추정하는 방법이다. 이 방법은 특징점의 수가 많아질수록 계산량이 복잡해질 뿐만 아니라 제약을 가하기가 어려워져 오류를 일으키기 쉽다. 고종호[3]는 하우스도프 거리(Hausdorff distance)를 기반으로 영상과 모델간의 유사도를 측정하여 대응영역을 검색하는 변형된 기법을 제안하였다. 이 방법은 영상의 세그멘테이션을 통해 영역을 분리하고, 분리된 영역내의 특징점을 그룹으로 하여 기준영상과 모델의 하우스도프 거리(Hausdorff distance)가 최소가 되는 대응을 변환관계로 선택하였다.

광류를 이용한 대응점 추정은 먼저 개략적인 중첩영역을 택하고 광류측정법(예를 들어 Lucas-Kanade법)을 이용하여 광류를 구한 다음, 이로부터 에피폴라 기하이론을 적용한 변환행렬을 구하는 방법이다[8]. 이 방법은 영상간의 대응을 변환행렬로 직접 구하므로 정확한 대응을 찾는 것이 가능하나 광류의 계산량이 많고, 노이즈의 영향을 받기 쉽다는 단점이 있다.

모자익을 위한 대응계산은 복수의 영상이 결합하여 하나의 대형 영상으로 합성된다는 점을 고려하여 볼 때 단순한 특징점 매칭 뿐만 아니라 전체 화소간의 유사관계를 고려함이 필요하다. 이러한 점을 고려하여 본 논문에서는 기존의 특징점 기반 대응점 검색 기법들이 사람의 개입에 의해 영상을 정합하거나, 형태정보만 이용하는 방법(하우스도프 디스턴스)을 사용 하는것과는 달리 형태정보와 칼라정보를 모두 사용하여 사람의 개입이 없이 자동으로 정확한 중첩영역을 검색하는 방법을 제안한다.

아래의 표 1은 대응 영역 검색 시 걸리는 시간과 검색 시 사용하는 정보 및 모델 영역 선택의 사람의 개입 여부를 알고리즘에 따라 분류한 결과이다.

표 1에서 m과 n은 각각 검색할 영상의 넓이(width)와 높이(height)를 나타낸다. 제안한 알고리듬의 시간 복잡도(time complexity)가 하우스도프(Hausdorff)검색방법과 같게 나오는 이유는 두 알고리듬 모두 이항검색을 통한 형태정보 정합이 지배항목이기 때문이다. 제안한 알고리듬은 형태정보 정합을 통해 후보영역을 선정한 후 그것에 대하여 색(질

표 1. 알고리즘에 따른 성능 분석표

	time complexity	형태, 질감 정보 사용	모델 영역 선택
제안한 알고리듬	$O(\log_2(m*n))$	형태정보 + (색)질감정보	자동 선택
Hausdorff	$O(\log_2(m*n))$	형태정보만 이용	수동 선택
단순 영역 비교	$O(m*n)$	색(질감)정보만 이용	수동 선택

감) 정보 정합을 실시하기 때문에 실제로는 하우스도프(Hausdorff)검색방법보다 약간의 시간이 더 필요하다. 그러나 제안한 알고리듬이 다른 검색방법과 비교하여 우수한 점은 형태정보와 질감정보를 모두 이용하여 매우 우수한 검색성공률을 보이면서 빠르게 검색을 수행할 수 있다는 점과 모델 선택을 자동으로 할 수 있다는 점이다.

본 논문에서는 위와 같은 사실들을 배경으로 하여 영상간의 중첩이 일어날 부분을 제한하여, 그 영역 안에서 특징점을 구하고, 다수의 특징점을 이용하여 Voronoi Distance(보로노이 거리)합이 최소가 되는 대응 계산 알고리즘을 제안한다. 이 방법의 특징은 다수 특징점들의 배열 구조가 유사한 대응영역을 찾아 대응후보 영역으로 등록한 뒤, 색상정보를 이용하여 최적 대응을 찾게 되므로 구현이 쉽고 처리 속도가 빠르다는 점이다. 또 화소 전체의 유사성을 고려하기 위해 먼저 구한 변환을 초기 모델로 하여 Levenberg-Marqudt 알고리즘에 의해 최적 변환 행렬을 구한다.

본 논문의 구성은 다음과 같다. 2절에서는 본 논문에서 제안한 보르노이 거리 기반 대응 영역 검색 방법 및 모델 생성을 위한 특징점 추출 방법에 대해 설명하고 3절에서는 변환행렬 계산 및 최종 영상 생성에 대해 기술하며 4절에서는 최종 영상 합성에 대해 기술하고 5절에서는 실험을 통하여 제안한 알고리즘의 효율성을 알아본다. 마지막으로 6절에서 결론을 맺고 향후 연구 과제에 대해서 언급한다.

2. 보르노이 거리 기반 대응 영역 검색

2.1 모자익 프로세스(Mosaic Process)와 중첩 제약

모자익 기술은 근본적으로 그래픽스에서 사용하는 영상워핑(Image Warping)이론과 컴퓨터 비전에서 사용하는 투영 변환(Perspective Transformation) 이론을 결합한 기술이다[11]. 모자이킹 프로세스는 크게 대응영역의 추정과 그것을 토대로 대응관계를 계산하는 과정, 대응관계로부터 영상을 합성하는 과정으로 나눌 수 있다. 이 중 대응 영역의 추정 프로세스는 초기 조건과 제약사항을 정함으로써 보다 빠르게 대응 관계를 계산하여 대응 영역을 검색할 수 있게 하는 중요한 기초 과정이다.

모자익 기술은 임의의 영상을 새로운 영상으로 변환시키기 때문에 각종 파라메터를 고려해야 되나 파라메터가 많으면 많을수록 대응영역의 검색이 어려워지고 변환행렬의 계산이 복잡하여 처리시간이 길어질 뿐만 아니라, 국소 극소점에 수렴할 가능성이 있다[10]. 이와 같은 이유로 본 논문에서는 합성할 영상의 그 중첩영역이 대략 영상의 2/3 영역 안에 있다고 가정하여 고려해야 할 파라메터의 수를 줄이고, 보로노이 평면(Voronoi Surface)을 이용하여 중첩영역에 대한 빠른 이항검색이 가능한 알고리즘을 제안한다. 이 제약은 비디오 카메라를 손에 들고 서서히 움직이면서 영상을 촬영할 때 타당한 제약이 된다.

본 논문에서 제안하는 영상 대응 관계 계산 과정은 그림 1과 같다. 비디오 카메라로부터 얻어진 영상들은 주로 병진 이동과 카메라의 XY축 방향의 회전에 의해 변화된 영상이며 그 중첩영역이 대략 영상의 2/3 영역안에 존재한다고 가정한다. 만일 XY축 회전이 미미하다면 주어진 영상의 2/3 영역 내에서 병진 이동에 대한 파라메터를 빠른 속도로 찾아낼 수 있다.

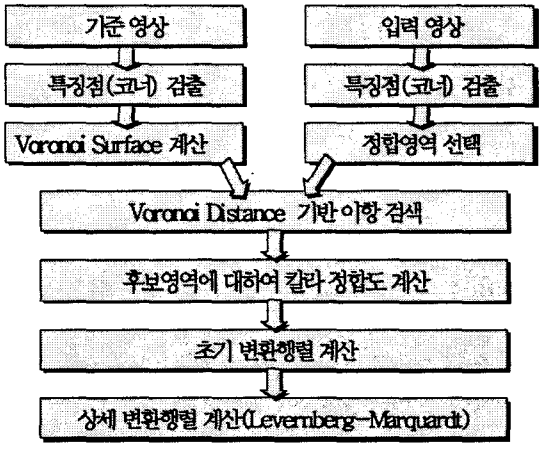

그림 1. 제안하는 영상 대응 관계 계산 과정

2.2 모델 영상 생성을 위한 특징점 추출

영상간의 대응관계는 눈에 띄는 특징점을 기준으로 주변의 밝기 변화를 비교하면 쉽게 찾을 수 있다. 본 논문에서는 특징점을 검출하기 위해 SUSAN 코너 검출기[5]를 사용하였다. 이 검출기는 밝기값 비교를 기반으로 하는 방법으로서 속도가 빠르고 신뢰

성이 뛰어나다. SUSAN 코너 검출기의 세부적인 내용은 다음과 같다.

영상 안에서 한 점 p(x,y)와 그 점을 중심으로 하는 마스크 윈도우를 W라고 정의한다. W안에서의 다른 한 점을 q(x,y)라 하고, I(p), I(q)가 각각 점 p, q에서의 밝기값을 나타낼 때 밝기값 유사도 측정은 다음과 같이 정의된다.

$$S(p,q)=\begin{cases} 1 & \text{if } |I(p)-I(q)|<t \\ 0 & otherwise \end{cases} \quad (1)$$

여기에서 t는 밝기값의 유사정도를 나타내는 문턱치이다. 점 p에서의 USAN(Univalue Segment Assimilating Nucleus)은 다음과 같이 계산된다.

$$N(p)=\sum_{w} S(p,q) \quad (2)$$

이 값이 마스크 윈도우 면적의 반보다 작으면 점 p를 특징점으로 본다. 따라서 이 값이 작을수록 점 p는 보다 확실한 SUSAN(Smallest USAN) 특징점이 된다.

그림 2는 SUSAN Corner Operator의 밝기에 대한 문턱치(t)를 변경시켰을 때 특징점의 획득 결과를 보인다. 원영상에 대하여 밝기 문턱치(Brighthness Threshold; BT)를 변경시킴에 따라서 특징점의 수

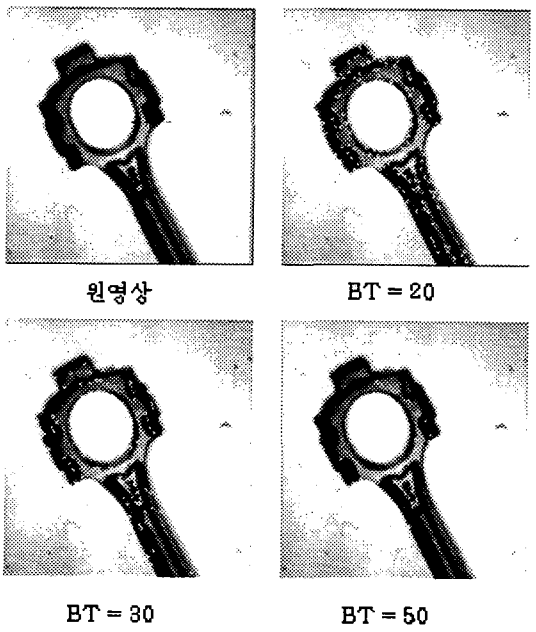

그림 2. SUSAN 코너 검출기의 밝기 문턱치 변환에 따른 결과

가 변하는 것을 볼 수 있으며, t가 낮을수록 중간 화소와 주변화소와의 명암차이의 수치가 낮아지므로 많은 특징점의 수를 얻을 수 있고, 반대의 경우 t가 높을수록 특징점의 수가 적어지는 것을 볼 수 있다. 본 논문에서 사용한 특징점은 추출된 코너점이 실세계 영상에서 3% 이하이므로 320*240영상에서 대략 2.6%에 해당하는 2000개를 최대값으로 설정하였다.

2.3 정합영역의 선택과 Voronoi Surface생성

기준영상과 입력영상에서 추출된 특징점들을 토대로 인접하는 영상간에 서로 중복되는 영역을 정한다. 비디오 카메라를 서서히 움직여 영상을 촬영한다고 가정했을 때 각 영상의 주변부에 중복되는 영역이 있다고 가정할 수 있다. 이 영역이 영상 정합에 있어서 신뢰도를 갖기 위해서는 특징점들이 골고루 분포되어 있어야 한다. 입력영상에 대해 특징점을 추출한 다음, 영상의 1/3이내의 주변부에 대해 특징점의 분포를 계산하고 그 값이 일정치 이상이면 이를 정합영역으로 선택한다.

기준영상에 대해 정합영역과의 매칭에 필요한 보로노이 평면(Voronoi Surface)를 생성한다. 여기서 정합영역은 보로노이 평면(Voronoi Surface)의 마스크로서 시작부터 끝까지 차례로 회선하여 보로노이 거리(Voronoi Distance)가 최소값이 되는 중첩영역을 찾는데 사용된다. 보로노이 평면(Voronoi Surface)는 보로노이 다이어그램(Voronoi Diagram)으로부터 영상을 특징점 중심의 거리로 표현하는 방법으로 구해진다. 보로노이 다이어그램(Voronoi Diagram)이란 한 평면상에 n개의 분리된 점들의 집합을 $P=\{p_1, p_2, p_3, \ldots, p_n\}$ 이라고 하고 이 때 두 점 p_i와 p_j $(i\neq j)$ 사이의 유클리드 거리를 $dist(p_i, p_j)$라고 하여 P의 보로노이 평면을 n개의 셀로 그림 3과 같이 나누는 것이다. P의 임의의 점 q에 대해서 만일, $dist(q, p_i)<dist(q, p_j)$를 만족하면 q는 p_i 셀에 포함된다. 평면상의 점들은 P에 있는 점들 중 가장 가까운 점이 속한 셀에 포함된다. 각 셀은 집합 P의 어떤 점에 대하여 가장 가까운 점들로 구성된다.

그림 3에서 보인 보로노이 다이어그램(Voronoi Diagram)에 대해 그림 4와 같이 집합 P의 점으로부

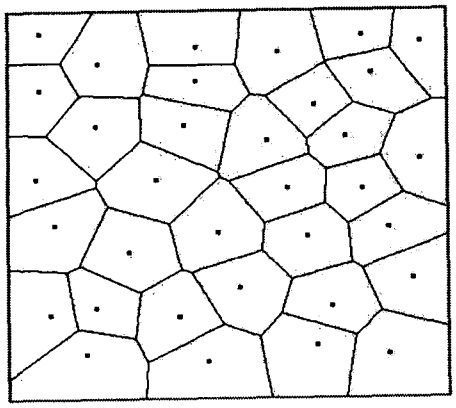

그림 3. 보로노이 평면, 검은 점이 SUSAN 알고리즘에 의해 구한 특징점 P

터의 거리를 높이로 생각하면 거리값이 0이 되는 점을 꼭지점으로 하는 원뿔들이 만들어진다. 원뿔의 꼭지점은 국소 극소점(local minima)이 되며 원뿔이 교차하는 지점은 2개의 국소 극소점들로부터 등거리에 있는 국소 극대점(local maxima)이 된다. 이 원뿔의 표면을 보로노이 평면(Voronoi Surface)이라고 한다. 보로노이 평면(Voronoi Surface)에 있는 한 점의 높이값은 특징점으로부터의 거리에 해당된다.

그림 4는 보로노이 평면(Voronoi Surface)를 옆에서 본 단면이다. 정합영역을 기준영상의 보로노이 평면(Voronoi Surface)위에 위치시키면 정합영역의 특징점에서 수선을 내려 맞닿는 점의 높이를 모두 합한 값이 보로노이 거리(Voronoi Distance)이다. 정합영역을 이동시켜 보로노이 평면(Voronoi Surface)의 특징점과 일치시키면 물리적으로 안정된 상태가 되고 보로노이 거리(Voronoi Distance)는 0이 된다.

본 논문에서는 SUSAN 코너 검출기에 의해 얻어진 기준영상의 코너를 토대로 보로노이 평면(Voronoi Surface)를 생성한다. 이 영상은 각 점에서 가장 가까운 특징점까지의 거리가 기록된 디지털 영상이 된다.

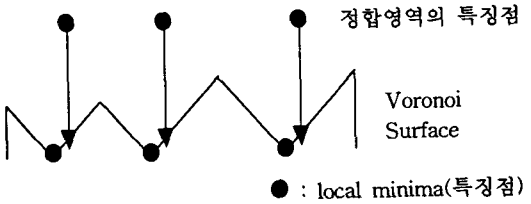

그림 4. 보로노이 평면의 단면과 정합영역의 특징점

2.4 Voronoi Distance 기반 이항 검색

우리가 1차적으로 찾고자 하는 것은 주어진 정합영역과 구조적으로 유사한 중첩 후보 영역들이다. 이것은 식(3)과 같이 주어진 문턱치내의 보로노이 거리(Voronoi Distance)를 갖는 변환 t를 찾는 것과 같다.

$$f(t) = V(A, t(B)) = \sum_{p_i \in t(B)} d(A, p_i) \le \tau \qquad (3)$$

여기서 $f(t)$는 보로노이 거리(Voronoi Distance), A는 기준영상의 보로노이 평면(Voronoi Surface), B는 입력영상의 정합영역, t는 변환행렬, $d(A, p_i)$는 A의 p_i에서의 높이값, τ는 문턱치를 나타낸다.

$f(t)$가 τ를 만족하는 영역들은 다수가 될 수 있는데, 만약 τ이 0이라면 정합영역과 정확히 일치하는 지점에서의 보로노이 거리(Voronoi Distance)가 0인 영역 1개만이 τ를 만족시킬 것이다. 그림 5는 정합영역과 정확히 일치하는 지점에서의 보로노이 거리(Voronoi Distance)를 계산한 것이다.

한편 $f(t) = V(A, t(B)) = \sum_{p_i \in t(B)} d(A, p_i)$ 는 t에 대한 선형방정식이므로 $f(t)$가 큰 값을 갖는다면 t가 근방의 다른 값 t'로 변화할 때 $f(t')$도 큰 값이 될 것을 예측할 수 있다.

이러한 사실을 바탕으로 식(4)를 유도할 수 있다.

$$|f(t) - f(t')| \le \sqrt{(t_x - t'_x)^2 + (t_y - t'_y)^2} \times N \qquad (4)$$

여기서 N은 정합영역의 특징점 개수를 나타내는 정수고, t에서 t'로 변화할 때 보로노이 거리(Voronoi Distance)의 차는 t와 t'의 유클리드 거리 $\sqrt{(t_x - t'_x)^2 + (t_y - t'_y)^2}$와 N의 곱보다 항상 적거나 같음을 알 수 있다.

그림 6의 왼쪽 그림은 t에 의하여 만들어지는 변

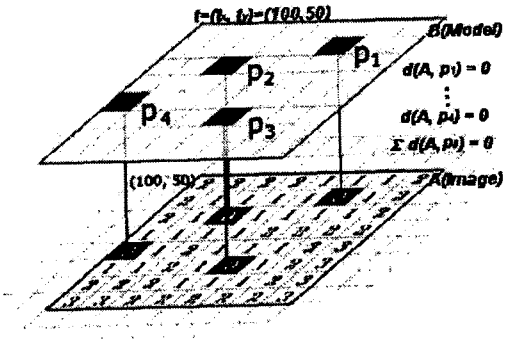

그림 5. 모델과 일치하는 영역에서의 Voronoi 거리 계산

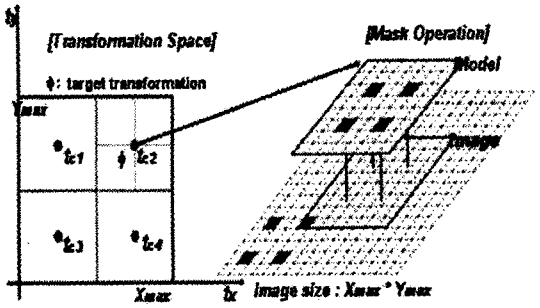

그림 6. 이항검색에 의한 중첩영역 검색의 예

환공간을 도식화한 것인데 보로노이 거리(Voronoi Distance)의 선형적인 특성을 이용하여 변환공간의 특정 영역내에서 찾고자 하는 변환 t_{target}이 존재할 가능성이 있는지 테스트해 볼 수 있다.

다음은 보로노이 거리(Voronoi Distance)를 기반으로 이항 검색을 수행하는 방법을 순서적으로 기술한 것이다.

단, 식(5)의 t_c는 변환공간의 (t_x^{high}, t_y^{high})에서 (t_x^{low}, t_y^{low})로 정의되는 사각영역의 중심점이다.

$$t_c = ((t_x^{low} + t_x^{high}), (t_y^{low} + t_y^{high})) \quad (5)$$

<알고리즘>

① 변환공간에서 모든 가능한 변환을 포함하는 사각형의 셀로부터 검색을 시작한다. 검색영역에 대한 리스트 전체를 포함하는 하나의 셀로 초기화한다.

② 검색리스트 내의 모든 셀을 스캔한다. 각 셀의 중심인 t_c에서 $V(A, t_c(B))$ 값을 계산한 후 $V(A, t(B)) \leq \tau$ 인 변환 t가 식(4)를 만족하는 경우에 해당 셀을 체크한다.

③ ②번에서 체크된 셀을 하위 차원의 검색리스트에 입력한다. 현재의 셀 크기가 충분히 작을 때까지 ②, ③번을 계속한다. 문턱치 τ가 셀의 크기에 비례하므로 초기에는 거의 모든 셀이 관심대상으로 체크된다.

위의 프로세스를 종료하면 최종적으로 기준영상의 중첩후보 영역들에 대한 변환으로 이루어진 리스트가 남게 된다.

2.5 중첩 후보 영역에 대한 칼라정합도 계산

보로노이 거리(Voronoi Distance)기반 이항 검색을 통해 남게 된 중첩 후보 영역들에 대하여 정합영역과의 칼라정합도를 계산한다. 칼라정합도는 식(6)의 비교영역의 픽셀간 에러에 의하여 계산된다.

$$e_{region} = \sum_{region} \{ | I_r^{red}(x,y) - I_m^{red}(x,y) | + | I_r^{green}(x,y) - I_m^{green}(x,y) | + | I_r^{blue}(x,y) - I_m^{blue}(x,y) | \} \quad (6)$$

I_r 은 기준영상에서 주어진 좌표의 칼라값을, I_m 은 정합영역에서의 칼라값을 나타내며 e_{region} 은 비교영역 전체의 에러를 합한 값이다.

후보 영역 중 에러 e_{region}이 가장 적은 영역이 바로 우리가 찾고자 하는 중첩영역이다.

3. 변환행렬 계산

3.1 초기변환 행렬의 계산

2차원 좌표계의 임의의 점을 $x=\{x, y, 1\}$이라고 할 때 2차원 투영변환에 의한 새로운 점의 좌표 x'는 다음과 같이 계산된다.

$$\boldsymbol{x}' = \boldsymbol{M}\boldsymbol{x} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & m_8 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (7)$$

여기서 m_0, m_1, m_3, m_4는 회전과 크기 요소, m_2, m_5는 이동요소, m_6, m_7는 비례변환요소, m_8는 1을 나타낸다. 이 식을 살펴보면 구하고자 하는 변수가 8개이므로 8개의 연립방정식이 필요하고, 4개의 대응 특징점에 의하여 두 영상간의 대응관계를 계산할 수 있다. 따라서, 검색된 중첩영역에서 4쌍의 대응 특징점을 선택하여야 한다. 이 때 대응 특징점은 중첩된 사각 영역의 8개의 꼭지점이다.

식(7)을 풀어쓰면 식(8),(9)를 얻을 수 있다.

$$x' = \frac{m_0 x + m_1 y + m_2}{m_6 x + m_7 y + 1} \quad (8)$$

$$y' = \frac{m_3 x + m_4 y + m_5}{m_6 x + m_7 y + 1} \quad (9)$$

초기 변환행렬 $\boldsymbol{M}$은 4쌍의 대응점을 통해 구할 수 있는 데 4쌍의 대응점에 대해 식(8), (9)를 적용하여 정리하면 식(10)을 얻을 수 있다.

3.2 최적 변환 행렬의 계산

(10)을 간단한 수치해석법에 의해 풀면 변환 행렬

$$\begin{bmatrix} x_1' \\ y_1' \\ x_2' \\ y_2' \\ x_3' \\ y_3' \\ x_4' \\ y_4' \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1'x_1 & -x_1'y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y_1'x_1 & -y_1'y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2'x_2 & -x_2'y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -y_2'x_2 & -y_2'y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3'x_3 & -x_3'y_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -y_3'x_3 & -y_3'y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4'x_4 & -x_4'y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -y_4'x_4 & -y_4'y_4 \end{bmatrix} \begin{bmatrix} m_0 \\ m_1 \\ m_2 \\ m_3 \\ m_4 \\ m_5 \\ m_6 \\ m_7 \end{bmatrix} \quad (10)$$

M을 구할 수 있다. 그러나 이렇게 계산된 초기 변환 행렬은 완전한 대응 특징 점을 토대로 구한 것이 아니기 때문에 오차를 갖는다. 따라서, 보다 많은 참조점을 이용하여 정확한 변환행렬을 구하는 것이 필요하다. 본 논문에서는 Levenberg-Marquadt 방법[1]을 이용하여 오차를 최소화한다. 이 방법은 중첩영역에서 최소가 되는 에러 ϵ를 구해 새로운 변환 행렬 M을 결정하는 방법으로 미지의 운동 요소 $\{m_0, \ldots, m_7\}$ 에 대한 ϵ의 편미분 값을 이용한다. 변환전의 영상과 변환후 영상간의 차이 e는 식(11)과 같이 정의한다. 식에서 $I(x_i, y_i)$는 pixel의 명도값(Intensity)을 나타낸다.

$$e = I'(x_i', y_i') - I(x_i, y_i) \quad (11)$$

최적 변환이란 식(11)의 값이 최소가 되는 변환행렬 M을 구하는 것이다. 여기에서 8개의 대응되는 운동요소의 미분은 식(8)과 (9)로부터 다음과 같이 구할 수 있다.

$$\left.\begin{aligned} \frac{\partial e}{\partial m_3} &= \frac{\partial e}{\partial m_0} = \frac{\partial x}{D}\frac{\partial I'}{\partial x'} \\ \frac{\partial e}{\partial m_4} &= \frac{\partial e}{\partial m_1} = \frac{\partial y}{D}\frac{\partial I'}{\partial x'} \\ \frac{\partial e}{\partial m_5} &= \frac{\partial e}{\partial m_2} = \frac{1}{D}\frac{\partial I'}{\partial x'} \\ \frac{\partial e}{\partial m_6} &= -\frac{x}{D}\left(x'\frac{\partial I'}{\partial x'} + y'\frac{\partial I'}{\partial y'}\right) \\ \frac{\partial e}{\partial m_7} &= -\frac{y}{D}\left(x'\frac{\partial I'}{\partial x'} + y'\frac{\partial I'}{\partial y'}\right) \end{aligned}\right\} \quad (12)$$

식 (12)에서 D는 식(8)의 분모를 나타낸다.

위의 편미분으로부터 Levenberg-Marquardt 법을 이용하여 식 (13)과 같이 헷세 행렬 A와 가중 기울기 벡터(Weighted Gradient Vector) b를 계산할 수 있다.

$$A = a_{kl} = \sum_i \frac{\partial e_i}{\partial m_k}\frac{\partial e_i}{\partial m_l}, \quad b = b_k = -\sum_i \frac{\partial e_i}{\partial m_k} \quad (13)$$

윗 식에서 i는 1부터 4까지의 대응점을, k와 l은 0부터 7까지의 운동 요소를 나타내는 첨자이다.

이 결과를 이용하여 Levenberg-Marquardt 알고리즘의 변위 계산식 $\triangle m = (A + \lambda I)^{-1} b$ 에 대입하여 운동 변수 $\triangle m$을 새로 계산하는 것이 가능하다. 여기에서 λ는 반복 계산에 대한 안정화 변수이다. 이 값은 식(11)에서 오차를 측정했을 때 이전 단계에서 최소화를 수행한 것보다 에러가 증가하면 그 값을 감소시키고, 감소하면 증가시켜 새로운 운동변수 $\triangle m$을 계산해 최적의 변환행렬 M을 구한다. 본 논문에서는 독립된 영역에 대해서만 위 알고리즘을 수행하여 최적의 변환행렬 M을 구하고, 결정된 변환행렬을 합성할 전체 영역에 최종적으로 적용시켜 계산시간을 줄였다.

4. 최종 영상의 합성

최종적으로 자연스러운 영상을 획득하기 위해서 영상을 합성하기 전에 각 영상에 블렌딩을 수행한다. 디지털 카메라나 비디오카메라로부터 얻어진 영상은 시선방향과 광원의 영향에 의해 영상들 사이에 밝기 차이가 나타나게 되며,이러한 이유로 재구성된 영상은 결합된 부분에 경계선(visible edge)을 가지게 된다. 본 논문에서는 이러한 경계선을 제거하기 위해 가중평균(weighted average)을 적용한 블렌딩 방법을 사용한다. 하지만, 영상을 단지 가중평균에 의한 방법만을 적용하면, 영상이 있는 부분과 없는 부분의 평균 계산 시 문제점이 생기게 되어, 오히려 영상이 끊어지는 현상이 발생하게 된다. 이러한 문제를 해결하기 위해, 변환될 영상에서 영상이 없는 부분의 값을 특별한 픽셀의 값을 이용하여 채운 다음, 가중 평균 계산 시, 이 값과 만나는 부분의 영상부분은 영상이 존재하는 영역의 픽셀값으로 대신한다.

식(14)와 식(15)는 눈에 보이는 경계선을 줄이기 위한 가중평균의 x와 y방향에 대한 블렌딩 함수를 나타낸다.

1) 이 방법은 비선형계의 최적치를 구하는 알고리즘이다.

$$w(x) = 1 - \left| \frac{2x}{width} - 1 \right| \quad w(y) = 1 - \left| \frac{2y}{height} - 1 \right| \tag{14}$$

$$w(x, y) = w(x)\, w(y) \tag{15}$$

위 식을 이용하여 A, B 두 영상을 합할 경우 새로 합성된 영상 (x, y)위치의 픽셀값은 식(16)으로 결정된다.

$$I'(x, y) = \frac{w_A(x, y) I_A(x, y) + w_B(x, y) I_B(x, y)}{w_A(x, y) + w_B(x, y)} \tag{16}$$

k개의 영상을 합성할 경우는 위의 관계로부터 아래 식을 사용하여 처리 할 수 있다.

$$I' = \frac{\sum_k w_k(x_i, y_i) I_k(x_i, y_i)}{\sum_k w_k(x_i, y_i)} \tag{17}$$

결국 식(17)을 이용하여 최종 영상 합성을 수행한다. 블렌딩을 수행한 후 이제 마지막으로 입력영상을 기준 영상에 대한 대응관계, 즉 2차원 투영변환 행렬을 적용하여 합성함으로써 하나의 대형영상이 나타나게 된다.

먼저 첫 번째 영상과 두 번째 영상간의 변환행렬 M_1을 구하고, 두 번째 영상과 세 번째 영상간의 변환행렬 M_2을 구하여 두 변환행렬의 곱을 최종 영상 합성 시에 적용하여 하나의 대형 영상을 생성한다.

5. 실험 결과

본 논문에서 제안한 알고리즘에 대해 개인용 컴퓨터(Pentium III-500, 128MB RAM, 그래픽 카드 AGP 16M)를 이용하여 실험 하였다. 실험에 이용한 영상은 SONY TRV 카메라를 사용하여 획득하였다. 실험은 실내 영상과 일반적인 실외 영상 그룹1, 밝기차가 크고 겹치는 영역이 적게 포함된 실외 영상 그룹2로 나누어 실험하였다.

테스트 영상들에 대한 실험 결과는 표 2와 같다. 여기에서 Threshold Factor는 식(3)과 식(4)의 문턱치를 증가시키는 정도를 나타낸다. 문턱치를 높일수록 유사도가 낮은 영역들이 후보로 포함된다. 표에서의 성공률은 실험단계에서 구해진 변환행렬을 오차 최소화 한 후에 각각의 테스트 영상(3장)들을 합성했을 때, 시각적으로 불연속적인 경계면이 발생하지 않고 안정적인 결과가 나왔을 때를 기준으로 정한 값이다. 나머지 수치값들은 각각의 알고리즘 적용 단계에서 처리에 걸린 시간(초)을 나타낸다. 표 2의 결과에 의하면 실내영상에 비해 실외영상의 성공률이 낮게 나타나고 처리에 걸린 시간도 더 길게 나타난다는 것과 겹치는 영역이 적게 포함된 영상도 성공률이 떨어진다는 것을 알 수 있다. 이는 밝기차가 심하게 나는 경우나 영상간의 겹치는 영역이 적을 경우에는 성공률이 낮게 나타난다는 것을 의미한다.

그림 7은 실내영상에 대하여 중첩영역을 검색한 결과이다. 그림에서 왼쪽 영상은 최초의 기준영상, 가운데 영상은 최초의 입력영상, 오른쪽 영상은 앞의 두 장의 영상이 합성한 결과로 생긴 기준영상에 대한 새로운 입력영상이다. 입력영상에서 선택된 최초의 정합영역(Model)과 구조적으로 유사한 후보 영역이 무작위로 아래 부분에 나열되어 있는 것을 볼 수 있다. 그리고 정합영역과 칼라정합도가 가장 큰 영역이 중첩영역으로 표시되어 나타나 있다. 여기에서 주목할 점은 실제 중첩영역의 근방에서도 보로노이 거리(Voronoi Distance) 값이 비슷하기 때문에 구조적 유사도만 가지고 정확한 중첩영역을 찾아낸다는 것은 무리가 있다. 따라서 구조적 유사도를 가지고 빠르게 후보영역들을 선택한 후에 칼라 정합도를 바탕으로 정확한 중첩영역을 정하였다. 그림 8은 실외영상을 가지고 중첩영역을 검색한 후 변환행렬을 계산하여 최종 합성한 결과이다. 나무들과 같은 복잡한 자연물에 대하여도 좋은 결과가 도출됨을 알 수 있다. 그림 9는 Threshold Factor를 4로 했을 때 실내 영상에 대한 결과를 보여주고 있다. 그림 10은 회전요소가 포함된 영상간의 중첩영역 검색 및 합성결과이다. 제

표 2. 입력 영상에 따른 실험 결과

	실내	실외1	실외2
Threshold factor	5	5	5
성공률	87%	81.1%	72.2%
특징점 검출	0.39 sec	0.39 sec	0.40 sec
Voronoi Surface생성	0.05 sec	0.05 sec	0.10 sec
모델 선택	0.71 sec	0.55 sec	0.65 sec
VD기반 이항검색	0.11 sec	0.09 sec	0.21 sec
칼라 매칭	0.17 sec	0.11 sec	0.74 sec
L-M 계산	2.69 sec	9.09 sec	10.08 sec

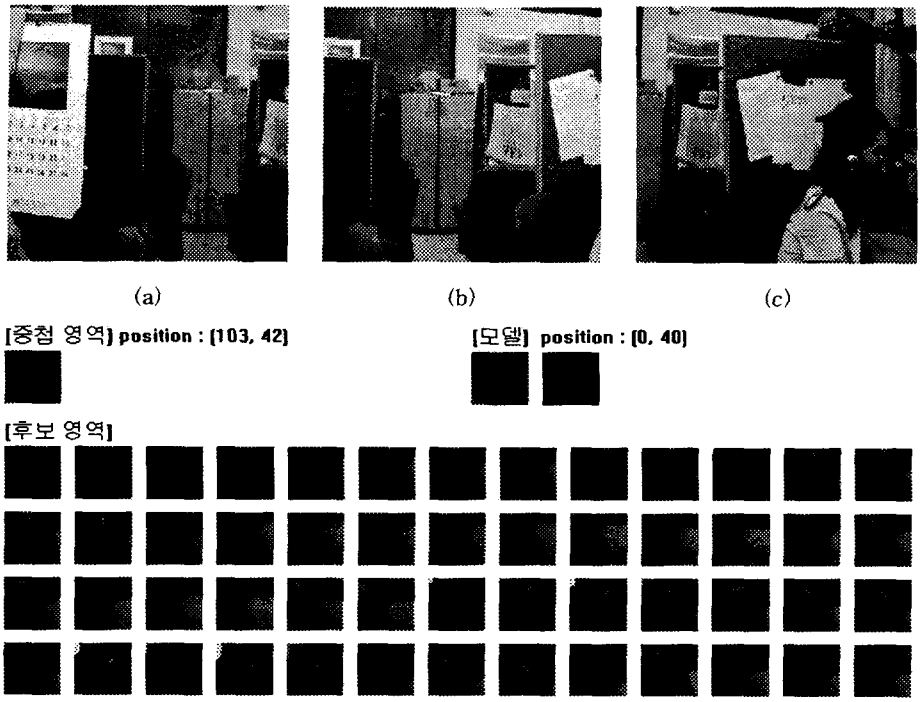

그림 7. 실내영상에 대하여 중첩영역을 검색한 결과
(a) 최초기준영상 (b) 최초입력영상 (c) 최종입력영상 (d)는 중첩영역 및 후보영역

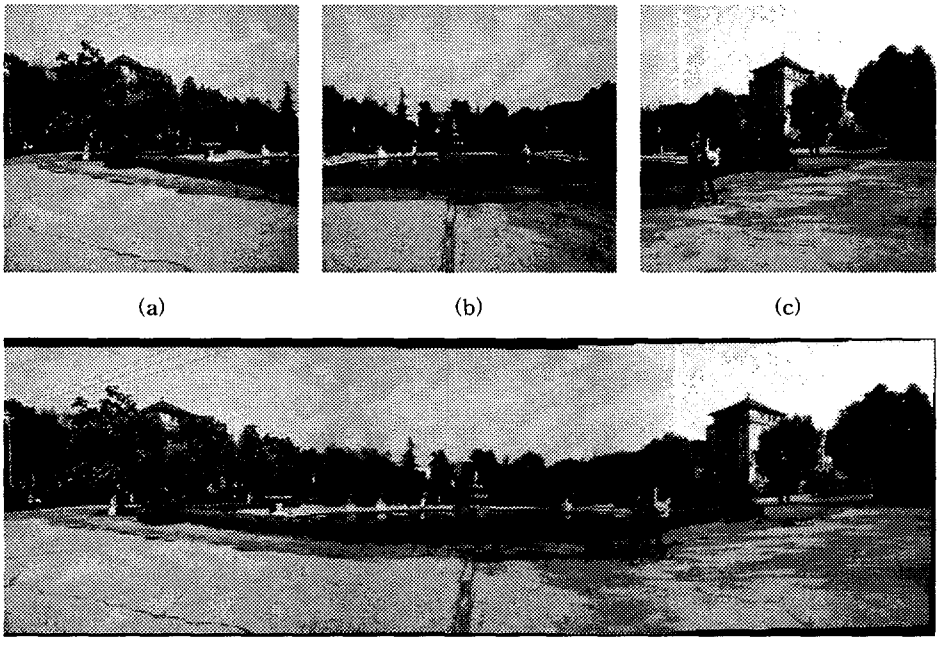

그림 8. 실외영상에 대한 검색과 합성 결과
(a),(b),(c)는 입력영상 (d)는 합성 결과

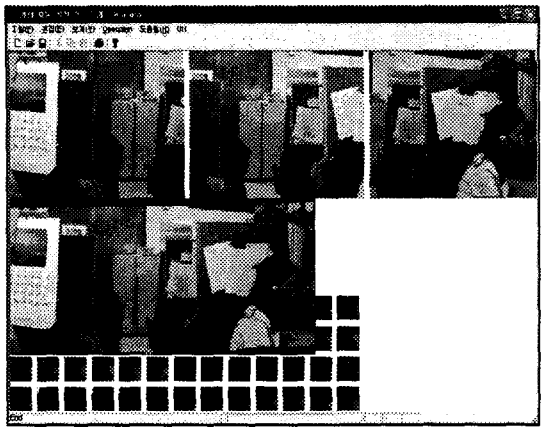

그림 9. 실내영상에 대한 검색과 합성 결과

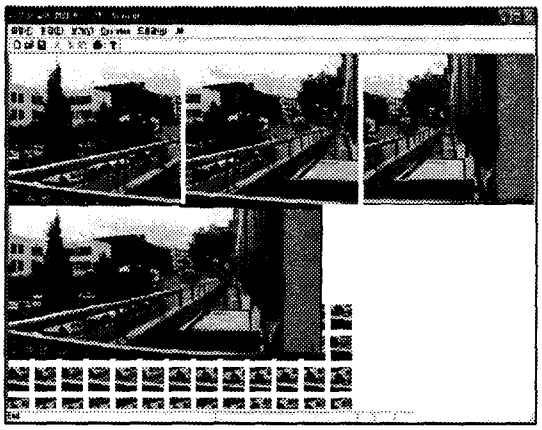

그림 10. 회전, 병진이동 요소를 모두 포함하는 영상간 중첩 영역 검색 및 합성 결과

안하는 알고리즘이 주로 병진 이동에 관하여 검색을 시도하나 회전에 대해서도 변화가 심하지 않는 범위 내에서 안정적인 결과를 나타냄을 알 수 있다.

6. 결론 및 향후 연구 계획

본 논문에서는 손에 든 비디오 카메라로 인접한 장면을 연결하여 촬영한 영상간의 중첩영역을 빠르고 정확하게 검색하는 방법에 대하여 기술하였다. 합성할 영상이 어느 정도 중복된 영역을 가지고 있다고 가정하여, 고려해야 할 파라메터의 수를 줄이고 보로노이 거리(Voronoi Distance)의 선형적인 특성에 기반한 이항 검색 기법을 사용하였다. 영역간의 구조적인 유사도와 칼라정합도를 차례로 이용함으로써 빠르면서 정확한 중첩영역 검색이 가능하도록 하였다.

실험을 통해 보로노이 거리(Voronoi Distance)정합법을 적용한 결과 실내영상은 물론 나무들과 같은 비교적 복잡한 자연물을 포함하고 있는 실외영상에 대해서도 좋은 결과가 도출됨을 알 수 있었다.

표 2의 실험결과를 보면 Levenberg-Marquardt법에 의해 오차를 최소화하는 시간이 상대적으로 높게 나타난다. 이것은 영상간의 오차 계산 시 행해지는 빈번한 양선형 보간법(Bilinear Interpolation)의 영향으로 일어나는 현상이다. 향후 연구 방향은 오차최소화 알고리즘을 개선하여 중첩영역 검색이후의 최종합성까지 소요되는 시간을 효과적으로 줄이는 방법에 대하여 연구할 계획이다.

참 고 문 헌

[1] Heung-Yeung Shum, R. Szeliski, "Panoramic image mosaicing", Technical Report MSR-TR-97-23, Microsoft Research, 1997.

[2] Daniel P. Huttenlocher, William J. Rucklidge, "A Multi-Resolution Technique for Comparing Images Using the Hausdorff Distance", TR 92-1321(Technical Report Department of Computer Science Cornell University) 1992. 12.

[3] 고종호, 이칠우, "특징기반 영상 모자이킹", IPIU2000(제 12회 영상 처리 및 이해에 관한 워크샵), pp275-280, 2000. 1.

[4] 이동휘, 최승현, 이칠우, "광역관찰 카메라 시스템을 위한 카메라의 대응관계 계산", 제 14회 신호처리 합동학술대회 논문집 pp415-418, 2001. 9.

[5] K, Wu and M.D. Levine, 2D Shape Segmentation: A New Approach, Pattern Recognition Letters, Vol. 17, No. 2, pp.133-140, 1996.

[6] R. Szeliski, "Video mosaics for virtual environments", IEEE Computer Graphics and Application, P22-30, March 1996.

[7] 김판규, "IDT (Incremental Delaunay Triangulation) 만들기", http:// pearl.cs.pusan.ac.kr/~pgkim/ idt.html, 1999.

[8] Naoki Chiba, Hiroshi Kano, Minoru Higashihara, Masashi Yasuda, and Osumi,

"Feature-Based Image Mosaicing", MVA'98 IAPR Workshop on Machine Vision Applications, pp5-10, 1998.

[9] Jezching Ton, Anil K.Jain, "Registering Landsat Images By Point Matching", IEEE Trans. on Geoscience and Remote Sensing, Vol.27, No.5, September, 1989.

[10] Bin Luo, E.R. Hancock, "Feature Matching with Procrustes Alignment and Graph Editing", IEEE Image Processing and its Applications, Conference Publicatioon No.465, 1999.

[11] Axel Pinz, Mafred Prantl, Harald Ganster "A Robust Affine Matching Algorithm Using an Exponentially Decreasing Distance Function", in Proc. CAIP'95, Computer Analysis of images and Patterns, V. Hlavac and R. Sara, Eds., vol.970 of LNCS, Springer, pp. 359-367, 1995.

이 칠 우

1986년 중앙대학교 전자공학과 졸업(학사)
1988년 중앙대학교 대학원 전자공학과(공학석사)
1992년 동경대학교 대학원 전자공학과(공학박사)
1992년~1995년 이미지 정보과학 연구소 수석 연구원겸 오사까대학 기초공학부 협력연구원
1995년 리츠메이칸 대학 특별초빙강사
1996~현재 전남대학교 전자컴퓨터정보통신 공학부 교수
관심분야 : 컴퓨터 비젼, 멀티미디어 데이터베이스, 컴퓨터그래픽스
E-mail : leecw@chonnam.ac.kr

정 민 영

1991년 숭실대학교 전자계산학과 졸업(학사)
1993년 숭실대학교 대학원 전자계산학과(공학 석사)
2002년 전남대학교 대학원 컴퓨터공학과(박사 수료)
1993년~1996년 숭실대학교 중앙전자계산소 연구원
1997년~1999년 광주여자대학교 정보전산원장
1996년~현재 광주여자대학교 멀티미디어학과 교수
관심분야 : 컴퓨터 비전, 멀티미디어 데이터베이스, 컴퓨터그래픽스, 소프트웨어 공학
E-mail : mychong@mail.kwu.ac.kr

배 기 태

1997년 호원대학교 전자계산학과 졸업(학사)
1999년 전남대학교 대학원 컴퓨터공학과(공학 석사)
2003년~현재 전남대학교 대학원 컴퓨터공학과(박사 수료)
관심분야 : 컴퓨터 비젼, 멀티미디어 데이터베이스, 패턴인식, 컴퓨터 그래픽스, 영상 처리
E-mail : bkt2002@empal.com

이 동 휘

2000년 전남대학교 컴퓨터공학학사
2002년 전남대학교 대학원 컴퓨터공학과 공학·석사
2002년~현재 (주)버추얼미디어 부설기술연구소 연구원
관심분야 : 컴퓨터 비젼, 영상 처리, 멀티미디어, 컴퓨터그래픽스

교 신 저 자

배 기 태 500-757 광주광역시 북구 용봉동 300 전남대학교 공과대학 컴퓨터공학과 공대 6호관 604호