Journal of Korea Multimedia Society Vol. 8, No. 9, September 2005(pp. 1195-1207)

기하 위상 정보를 이용한 고속 영상 정합 기법

배기태[†], 정민영[††], 오재용[†††], 이칠우[††††]

요 약

본 논문에서는 기하 위상 정보를 이용한 영상 정합 기법에 관해 기술한다. 기존의 특징 점 기반 대응점 검색 기법들이 주로 평행 이동 형 형태나 약간의 회전 및 상하 이동요소가 포함된 영상만을 처리 하는 것과 달리 큰 회전이나 상하이동요소가 포함된 입력영상들의 정확한 중첩영역을 자동으로 빠르게 검색하는 방법을 제안한다. 제안한 방법은 영상내의 특징 점들의 기하 위상 정보와 특징점 사이의 거리 정보를 가지는 DT(Distance Transform) map을 이용하여 정합 위치를 찾기 위한 계산 량을 크게 감축시키고 특히 회전이나 상하 이동요소가 포함된 영상들 간의 정합영역을 검색하는데 있어서 좋은 성능을 나타낸다.

High-Speed Image Matching Method Using Geometry-Phase Information

Ki-Tae Bae[†], Min-Yeong Chong[††], Jae-Yong Oh[†††], Chil-Woo Lee[††††]

ABSTRACT

In this paper, we describe image matching techniques which is automatically retrieving the exact matching area using geometry-phase information. We proposed a Matching Method which is rapidly estimating the correspondent points between adjacent images that included big-rotation and top-bottom movement element. It is a method that reduce computaion quantity to be required to find an exact correspondent position using geometry-phase information of extracted points in images and DT map which set the distance value among feature points and other points on the basis of each feature point of a image. The proposed method shows good performance especially in the part to search a exact correspondent position between adjacent images that included big-rotation and top-bottom movement element.

Key words: Image Matching(영상합성), Vornnoi Surface(보로노이 평면), Distance Transform

1. 서 론

영상 모자익 기술은 중첩된 여러 장의 영상에 대해 공통 부분을 찾아내어 하나의 고해상도 영상으로 통합 시킴은 물론 계산에 의해 시점을 달리하여 영상을 표현할 수 있다. 이 기술을 이용하면 그래픽 모델을 사용하지 않고도 복잡한 장면들을 간단하게 생성할 수 있기 때문에 영상 생성이 필요한 각종 응용 시스템에 많이 활용 될 수 있다. 특히 실영상을 입체적으로 조작할 수 있게 되면서 이제는 컴퓨터 그래픽

※ 교신저자(Corresponding Author) : 배기태, 주소 : 광주광역시 북구 용봉동 300(500-757), 전화 : 062)530-0258, FAX : 062)530-1759, E-mail : bkt2002@image.chonnam.ac.kr
접수일 : 2005년 3월 9일, 완료일 : 2005년 4월 26일
[†] 준회원, 전남대학교 대학원 컴퓨터공학과 박사과정
[††] 정회원, 광주여자대학교 멀티미디어학과 교수
(E-mail : mychong@mail.kwu.ac.kr)
[†††] 전남대학교 대학원 컴퓨터공학과 박사과정
(E-mail : ojyong@image.chonnam.ac.kr)
[††††] 종신회원, 전남대학교 전자컴퓨터정보통신 공학부 교수
(E-mail : leecw@chonnam.ac.kr)
※ 본 연구는 전남대학교 "고품질 전기 전자 부품 및 시스템 연구센터"의 연구비 지원에 의해 수행 되었음.

스의 페인팅 시스템, 가상현실 시스템, 영화제작 등 많은 분야에서 활용되고 있다[1].

모자익 알고리즘에 있어서 중요한 요소기술은 두 영상을 접합할 중첩영역을 찾는 기술(영상 정합 과정)과 중첩영역을 기준으로 여러 영상을 불연속점 또는 왜곡 없이 매끄럽게 합성하는 기술이다. 복수의 영상간의 대응관계를 찾는 것은 스테레오 정합, 트랙킹, 움직임 추정, 영상교정, 3차원 스테레오 기하에 의한 3차원 정보 추출 등 컴퓨터 비젼 분야에서 중요한 부분으로 인식되어 왔다.

두 영상의 중첩영역을 찾는 기술들은 영상간의 상관을 이용하거나[2,12], 또는 광류[3]를 이용하는 방법, 특징점간의 거리 제약[4,13]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. 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[3]. 이 방법은 영상간의 대응을 변환행렬로 직접 구하므로 정확한 대응을 찾는 것이 가능하나 광류의 계산 량이 많고, 노이즈의 영향을 받기 쉽다는 단점이 있다. 특징 점간의 거리 제약을 이용하는 방법[4]은 찾아낸 특징 점들이 영상 간에 서로 임의의 대응을 갖는다고 가정하고 그 중 거리 오차가 가장 적은 결합을 영상 대응으로 추정하는 방법이다. 이 방법은 영상의 에지와 같은 저수준의 특징을 이용하는 방법으로 HCM (Hierarchical Chamfer Matching)[5], HD(Hausdorff Distance)[2,6]등이 그 대표적인 예라고 할 수 있다. HCM방법은 물체의 에지로부터 특징 점을 얻은 후 특징점 사이의 유사도에 의해 물체 정합을 수행하는 방법이다. HD정합 방법은 에지로부터 특징 점을 구하지 않고 전체 에지 점집합들 사이의 HD에 의해 물체의 유사도를 판단하는 방법이다. 이 방법은 이진 영상내의 모든 점에 대한 거리 계산 과정을 수행해야 하므로 특징 점의 수가 많아질수록 계산 량이 복잡해질 뿐만 아니라 제약을 가하기가 어려워져 오류를 일으키기 쉽다. 이러한 단점을 해결하기 위해 HD 정합 방법은 에지 점집합들 사이의 거리 계산을 효과적이고 빠르게 처리할 수 있는 DT(Distance Transform) map[6,7,11]을 사용하였다.

위에서 언급한 방법들은 정확한 정합 위치를 찾는데 많은 시간이 걸릴 뿐만 아니라 회전이나 상하 이동을 고려하지 않은 좌우 이동 형태의 영상만을 처리하기 때문에 확장성이 떨어지는 단점이 있다. 또한 카메라 회전각이 상하 좌우로 크게 차이가 나는 상태에서 획득한 영상들에 대해서 정합 영역을 찾는 것은 더더욱 힘들다.

본 논문에서는 위와 같은 영상 정합에 있어 어려운 점을 고려하여 회전이나 상하이동요소가 포함된 입력영상들을 사람이 개입하지 않고 자동으로 정확한 중첩영역을 빠르게 검색하는 방법을 제안한다. 제안한 방법의 특징은 영상간의 중첩이 일어날 부분을 제한하여 그 영역 안에서 특징 점을 구하고, 영역 내의 특징 점들의 기하 위상 정보와 특징점 사이의 거리 정보를 가지는 DT map을 이용하여 특징 점들의 배열 구조가 유사한 대응영역을 찾아 최적의 대응을 찾게 되므로 구현이 쉽고 처리 속도가 빠르다는 점이다.

본 논문의 구성은 다음과 같다. 2절에서는 본 논문에서 제안한 기하 위상 정보 기반 영상 정합에 대해 설명하고 3절에서는 영상 합성과정 에 대해 기술하며 4절에서는 실험을 통하여 제안한 알고리즘의 효율성을 알아보고 결론을 맺고 향후 연구 과제에 대해서 언급한다.

2. 기하 위상 정보 기반 영상 정합

2.1 영상정합 과정

모자익 프로세스는 크게 대응영역의 추정과 그것을 토대로 대응관계를 계산하는 과정(영상 정합), 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[8]. 이와 같은 이유로 본 논문에서는 카메라로부터 얻어진 영상들은 주로 병진 이동과 카메라의 XY축 방향의 회전에 의해 변화된 영상이며 그 중첩

영역이 대략 영상의 2/3 영역 안에 존재한다고 가정하였다.

본 논문에서 제안하는 기하 위상 정보 기반 영상 정합 과정은 그림 1과 같다.

그림 1. 기하 위상 정보 기반 영상 정합 과정

2.2 Harris Corner Detector를 이용한 특징 추출 및 모델 선택

영상간의 대응관계는 눈에 띄는 특징 점을 기준으로 주변의 밝기 변화를 비교하면 쉽게 찾을 수 있다. 본 논문에서는 특징 점을 검출하기 위해 Harris 코너 검출기[9]를 사용하였다. 이 검출기는 밝기 값 비교를 기반으로 하는 방법으로서 영상의 각 점의 명도에 대해 x와 y의 2차 편미분 값을 구성요소로 하는 헷세 행렬의 행렬식 값을 계산하여 그 극대점을 찾아 특징점을 검출하는 방법으로 viewpoint 변화와 조명 변화에 불변하는 특징이 있다. Harris 코너 검출기의 세부적인 내용은 다음과 같다.

영상 안에서 한 점을 $I(x,y)$라고 정의했을 때 x와 y의 2차 편미분 값을 구하여 (1)과 같은 헷세 행렬을 생성한다.

$$M=\begin{bmatrix} \left(\frac{\partial I}{\partial x}\right)^2 & \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) \\ \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) & \left(\frac{\partial I}{\partial y}\right)^2 \end{bmatrix} \tag{1}$$

생성된 헷세행렬 M을 이용하여 영상의 각 점마다 (2)에서 정의한 행렬식 R을 계산한다. 정해진 윈도우 영역 내에서 R값을 비교한 후 가장 큰 R값을 가진 점을 특징 점으로 선택한다.

$$R=\det M-k(traceM)^2 \tag{2}$$

(3)은 (2)에서 정의한 식을 행렬식의 고유 값으로 나타낸 식이다. 만일 영상안의 어떤 점에서 행렬 M의 2개의 고유값 λ_1과 λ_2의 값이 크게 나왔다면 이는 그레이 레벨의 변화가 심하게 일어난 곳을 의미하며 이 점이 특징 점으로 선택된다.

$$C=\begin{bmatrix} \Sigma I_x^2 & \Sigma I_x I_y \\ \Sigma I_x I_y & \Sigma I_y^2 \end{bmatrix}=\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{3}$$

추출된 특징 점들의 고유 값들의 집합을 $L=\{\lambda_1,\lambda_2,\cdots\cdots\lambda_k,\cdots\cdots\lambda_n\}$라고 하고 모델영역의 특징점들의 집합을 $M=\{p_1,p_2,\cdots\cdots p_k,\cdots\cdots p_n\}$이라고 정의했을 때, 본 논문에서는 모델영역에서 선택된 특징점이 검색영역에서 누락되지 않게 하기 위해 모델영역에서 구해진 특징 점들을 고유값 순으로 정렬하여 상위 6개를 최종 3쌍의 후보로 선택한다. 이는 고유값이 큰 값들이 영상에서 두드러진 코너값일 확률이 높기 때문에 두 영상에 모두 코너점으로 추출될 가능성이 높다는것을 의미한다.

2.3 모델 및 기준 영상의 후보 선택 및 기하 위상 정보 계산

모델 영상에서 선택된 3쌍의 후보와 기준 영상의 특징 점들 사이의 유클리드 거리를 비교하여 일정 거리 안에 있는 쌍들을 후보로 선택한다. 앞 절에서 구해진 후보 특징 점들의 집합을 $A=\{a_1,a_2,a_3,a_4,a_5,a_6\}$ 라고 하고 각각의 후보 쌍들의 유클리드 거리를 계산한 집합을 $U=\{u_1,u_2,u_3\}$라고 정의한다. 이때 두 점을 a_i와 a_j $(i\neq j)$, 두점 사이의 유클리드 거리를 $dist(a_i,a_j)$라고 했을 때, (4)와같이 세 쌍의 후보에 대한 유클리드 거리를 계산한다. (4)에서 x_t는 a_i의 x좌표, y_t는 a_i의 y좌표, $x_{t'}$는 a_j의 x좌표, $y_{t'}$는 a_j의 y좌표를 의미한다.

$$u_i=dist(a_i,a_j)=\sqrt{(x_t-x_{t'})^2+(y_t-y_{t'})^2} \tag{4}$$

(4)에서 구해진 유클리드 거리를 이용하여 세 후

보 쌍 각각에 대해 일정 범위 안에 있는 검색영상의 특징 쌍들을 후보로 선택한다. 검색 영상 특징 점들의 유클리드 거리를 $V=\{v_1,...,v_q\}$(q는 전체 후보 쌍의 수) 라고 정의하고 유클리드 거리 비교 후 후보들의 집합을 $C=\{c_1,...,c_n\}(c_i=\{a_1,...,a_m\})$ 라고 정의했을 때 c_i는 (5)을 만족하는 특징 쌍들을 후보로 선택한다. ϵ값은 후보를 선택하는 기준 값으로 ϵ값이 크면 후보의 수가 많아져서 검색 속도가 느려지고 ϵ값이 작으면 후보의 수가 작아져서 검색 속도는 빨라지는 반면 영상의 회전이나 변형이 있는 경우 정확한 대응 후보쌍이 누락될 가능성이 있다. 따라서 본 논문에서는 실험을 통해 가장 적절한 값으로 ε값을 3으로 설정 하였다.

$$c_i=\begin{Bmatrix} O & \text{if}(u_i-\epsilon<v_j<u_i+\epsilon) \\ X & otherwise \end{Bmatrix} \quad (5)$$
$$(0\le i\le 2, 0\le j\le q)(O=select, X=not\ select)$$

그림 2는 모델 영상에서 선택된 1쌍의 후보와 (5)의 범위 내에 있는 유클리드 거리를 가지는 후보들이 검색영상에서 검출된 모습이다.

회전된 영상간의 정확한 대응점 검색을 위해서는 모델영역과 검색영역 간의 정확한 위상각의 차이를 알아야 한다. 따라서 검색영역에서 선택된 후보 쌍들의 위상각과 모델 영역에서 선택되어진 3쌍의 후보들과의 위상각의 차이를 구한다. 모델후보들의 위상각을 저장한 집합을 $R=\{r_1,r_2,r_3\}$, 검색영역에서 선택된 후보들의 위상각을 저장한 집합을 $Q=\{q_i,q_j,q_k\}(q_i=\{\theta_1,...,\theta_l\},q_j=\{\theta_1,...,\theta_m\},q_k=\{\theta_1,...,\theta_n\})$, 선택된 후보들과 모델후보의 위상각의 차이를 저장한 집합을 $T=\{t_i,t_j,t_k\}(t_i=\{\theta_1,...,\theta_l\},t_j=\{\theta_1,...,\theta_m\},t_k=\{\theta_1,...,\theta_n\})$ 라고 정의했을 때 (6)은 두 모델간의 위상각의 차이를 구한 식을 나타낸다.

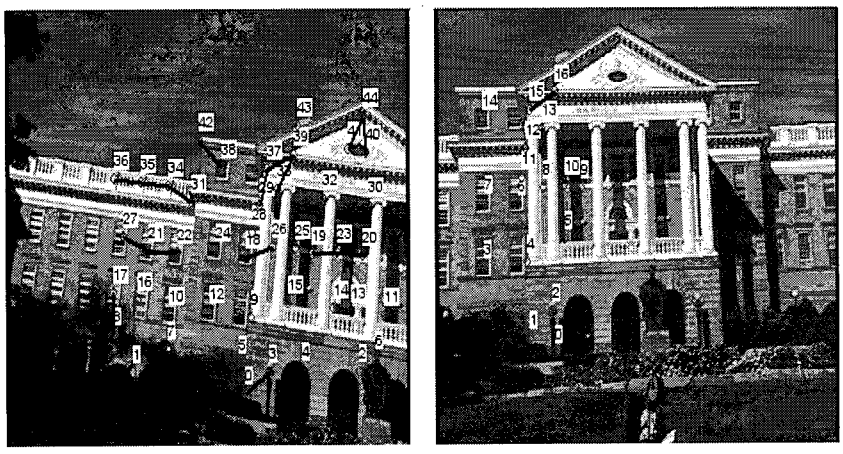
$$t_i=(q_i-r_1),t_j=(q_j-r_2),t_k=(q_k-r_3) \quad (6)$$
$$(1\le i\le l,1\le j\le m,1\le k\le n)$$

(6)에서 구해진 위상각 정보를 이용하여 모델 영역에 있는 특징 점들을 회전이동 시켜준다. 이는 다음단계의 보로노이 거리 정합 시 회전된 모델 영역을 DT map에 겹친 후 보로노이 거리를 구하는데 이용된다.

모델 영역의 집합을 $M=\{p_1,...,p_m\}$, 모델 영역의 한 점의 좌표를 $p=(x_{org},y_{org})$,회전 이동된 후의 집합을 $M=\{p'_1,...p'_m\}$, 한 점의 좌표를 $p'=(x_{new},y_{new})$, (5)에서 선택된 기준 후보들의 좌표 집합을 $S=\{s_1,s_2,s_3\}$, 한 점의 좌표를 $s=(s_x,s_y)$라고 정의 했을 때, 모델영역의 한 점은 (7)과 같은 형태로 미리 구해진 위상각에 따라 회전 이동 된다. (7)에서 θ값은 (6)에서 구해진 T값을 적용한다. (8)은 모델영역이 기준좌표 S를 중심으로 (6)에서 구해진 T에 의해 회전 이동 된 것을 나타낸다.

$$\begin{pmatrix} x_{new} \\ y_{new} \end{pmatrix}=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x_{org}+s_x \\ y_{org}+s_y \end{pmatrix}+\begin{pmatrix} s_x \\ s_y \end{pmatrix} \quad (7)$$

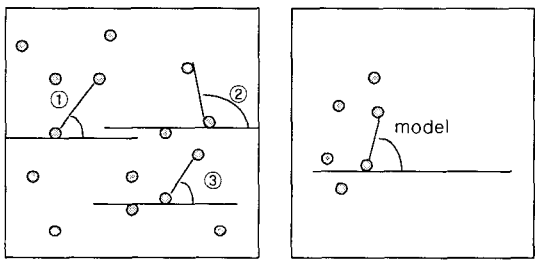
$$M=\begin{pmatrix} \cos t_i & -\sin t_i \\ \sin t_i & \cos t_i \end{pmatrix}(M+s_i)+(s_i),\ (1\le i\le 3) \quad (8)$$



(a) (b)

그림 2. 검색영역(a)에서 검출된 후보 쌍들과 모델영역(b)에서의 기준 후보 쌍

최종적으로 모델영역을 구해진 검색영역과의 위상각의 차이에 따라 회전 이동시킨 후 결과를 저장한 집합을 $W=\{w_1,...,w_n\}$라고 정의 했을 때, w_i는 모델영역을 (7)에서 구한 c_i의 각각의 후보 값에 대해 (6)에서 구한 위상각 집합 t_i만큼 i번째 모델후보의 기준좌표를 중심으로 회전이동 시킨 결과를 나타낸다. 그림 3은 검출된 후보들과 모델영역의 기준후보들의 위상각을 계산하는 과정을 나타낸 것이다. 그림 3의 (b)는 모델영역의 특징점을 나타내고 이들은 각각 계산된 위상각에 따라서 회전 이동된 후 DT map에 겹쳐지게 된다.



(a) (b)

θ_1 =① - model
θ_2 =② - model
θ_3 =③ - model

(c)

그림 3. (a) 검출된 후보 쌍들의 위상각, (b) 모델영역의 기준 후보 쌍의 위상각, (c) 후보와의 위상각 차이.

2.4 DT map 생성 및 보로노이 평면 생성

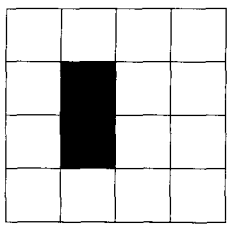
기준영상에 대해 모델영역과의 매칭에 필요한 DT(Distance Transform) map을 생성한다. 여기서 모델영역은 앞 단계에서 구해진 각각의 후보들의 위상각을 기준으로 회전이동 된 후 DT map상의 보로노이 거리가 최소값이 되는 중첩영역을 찾는데 사용된다. DT map은 보로노이 다이어그램(Voronoi Diagram)으로부터 영상을 특징 점 중심의 거리로 표현하는 방법으로 구해진다. 모델 영역은 DT map의 마스크로서 시작부터 끝까지 차례로 회선하여 보로노이 거리(Voronoi Distance)의 합이 최소가 되는 정합영역을 찾는데 사용된다. 따라서 기준 영상에 대해 모델 영역과의 매칭에 필요한 보르노이 평면을 생성해야 한다.

보로노이 평면은 기준 영상의 각 요소 (x,y)에서 가장 가까운 특징 점과의 거리정보를 정의하고 있는 2D평면(2차원 배열)을 말하는데, 각 요소의 거리 정보를 보로노이 거리라 한다. 보로노이 평면의 각 요소의 위치는 기준 영상의 위치와 동일하며 다만 각 요소의 값이 색깔 값이 아니라 가장 가까운 특징 점과의 거리 값인 것이 다른 점이라고 할 수 있다.

보로노이 평면을 생성하기 위해서는 보로노이 평면의 각 요소에 대한 보로노이 거리를 계산해야 한다. 기본적인 보로노이 거리 계산 방법은 보로노이 평면 각 요소 (x,y)에서 기준 영상의 각 특징 점 좌표와의 x축 거리와 y축 거리 중 큰 것을 구한 다음 이들 중에서 가장 작은 값으로 정하면 된다. 이렇게 하면 특징점 주변 대각 위치 거리 값이 x위치나 y위치의 값과 같게 되므로 그림 4와 같은 결과를 보인다.

그림 5는 실제 영상을 외곽선을 검출한 후 구해진 에지점 들을 기준으로 DT map을 생성한 그림이다. 특징점에서 멀어질수록 밝아지고 특징점에 가까울수록 어두워지는 것을 알수 있다.

그림 4에서 보인 보로노이 다이어그램에 대해 그림 6과 같이 집합 P의 점으로부터의 거리를 높이로 생각하면 거리 값이 0이 되는 점을 꼭지 점으로 하는 원뿔들이 만들어진다. 원뿔의 꼭지 점은 국소 극소점(local minima)이 되며 원뿔이 교차하는 지점은 2개의 국소 극소점들로부터 등거리에 있는 국소극대점(local maxima)이 된다. 이 원뿔의 표면을 보로노이 평면(Voronoi Surface) 이라고 한다. 보로노이 평면



2	1	2	3
1	0	1	2
1	0	1	2
2	1	2	3

그림 4. Distance Transform Map 검은 부분이 Harris 알고리즘에 의해 구한 특징점 P

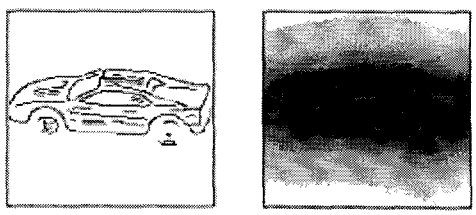


그림 5. 실제 영상에 적용한 후의 DT map

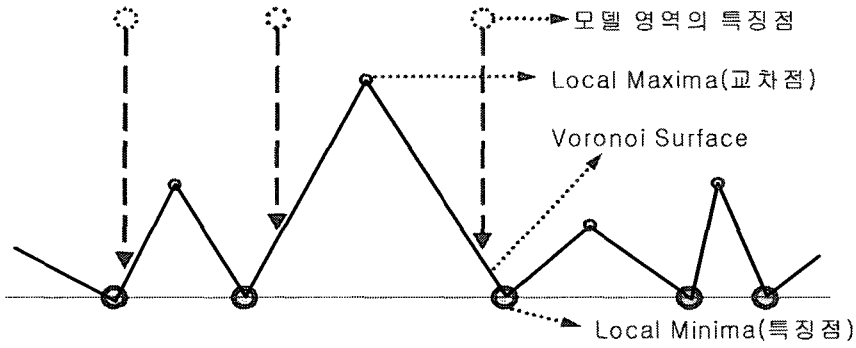


그림 6. 보로노이 평면의 단면과 정합영역의 특징점

에 있는 한 점의 높이 값은 특징 점으로부터의 거리에 해당된다.

그림 6은 보로노이 평면을 옆에서 본 단면이다. 각각 회전 이동된 모델영역을 기준영상의 보로노이 평면위에 위치시키면 모델영역의 특징 점에서 수선을 내려 맞닿는 점의 높이를 모두 합한 값이 보로노이 거리이다. 모델영역을 일치하는 위상각에 맞게 회전 이동시켜 보로노이 평면의 특징 점과 일치시키면 물리적으로 안정된 상태가 되고 보로노이 거리는 0이 된다.

본 논문에서는 Harris 코너 검출기에 의해 얻어진 기준영상의 코너를 토대로 보로노이 평면을 생성한다. 이 영상은 각 점에서 가장 가까운 특징 점까지의 거리가 기록된 디지털 영상이 된다.

2.5 기하위상 정보를 이용한 최종 후보 선택

우리가 찾고자 하는 것은 주어진 모델영역과 구조적으로 유사한 중첩 영역이다. (9)는 보로노이 거리를 계산하는 변환 t를 나타내고 있다.

$$f(t) = V(A;t(B)) = \sum_{p_i \in t(B)} d(A,p_i) \tag{9}$$

여기서 $f(t)$는 보로노이 거리값, A는 기준영상의 보로노이 평면, B는 입력영상의 모델영역, t는 변환행렬, $d(A,p_i)$는 A의 p_i에서의 높이 값을 나타낸다.

모델영역과 정확히 일치하는 지점에서의 보로노이 거리는 0을 나타내고 그림 7은 모델영역과 정확히 일치하는 지점에서의 보로노이 거리를 계산한 것이다. 그림 8은 2차원 평면상에서 보르노이 거리를 계산하는 방법을 나타내고 있다.

세 후보 쌍들의 위상정보를 기준으로 회전 이동된

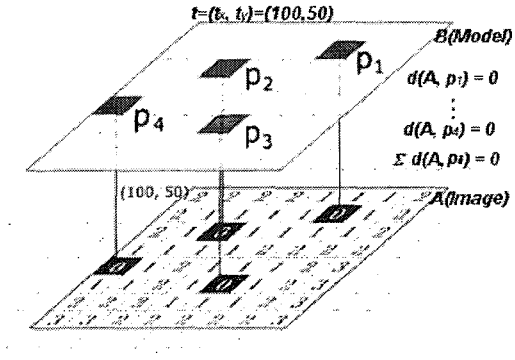


그림 7. 모델과 일치하는 영역에서의 Voronoi 거리 계산

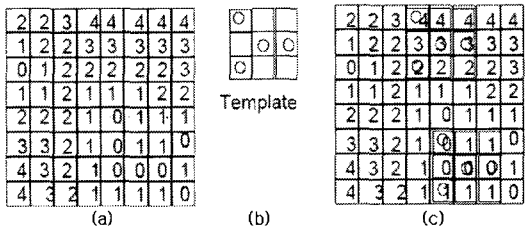


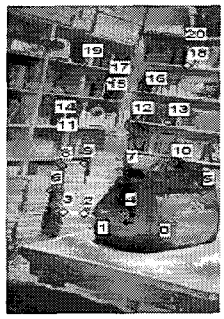
그림 8. 보르노이 거리 계산(2차원 평면상)

각각의 모델영역의 $f(t)$값을 구한 후 가장 값이 작은 모델영역을 최종 후보로 선택한다. (10)에서 M은 가장 작은 값을 가진 모델영역을 가리키며, n은 후보수를 나타낸다.

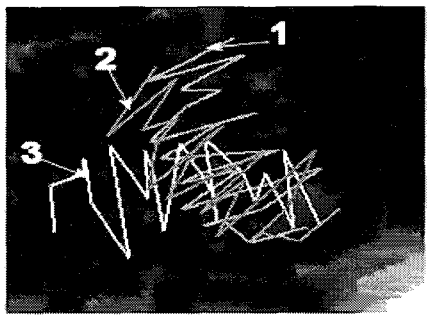
$$M = \min(f(t_1), f(t_2), \ldots f(t_n)) \tag{10}$$

그림 9는 세 후보 쌍의 보로노이 거리를 계산한 결과를 보여주고 있다. 그림 9의 (d)를 보면 후보들의 보로노이 거리 값이 나와 있는데, 수치를 보면 첫 번째 후보가 가장 작은 값을 나타내고 있는 걸 볼 수 있다. 이는 첫 번째 후보 모델영역의 특징 점들이 검

(a)



(b)



(c)

번호	브르노이 거리값
1	20
2	210
3	320

(d)

그림 9. 세 후보 쌍의 보로노이 거리 검색: (a) 검색영역상의 매칭 결과, (b) 모델(특징점), (c) DT map상의 매칭결과, (d) 각 후보의 보로노이 거리값.

색영역의 가장 비슷한 위치에 겹쳐졌다는 것을 의미하며, 이 모델영역이 최종 후보로 선택된다. 그림 9의 (c)는 검색영역의 특징 점을 기준으로 DT map을 생성한 결과이다.

3. 영상 합성 과정

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.

$$\boldsymbol{x}' = \boldsymbol{M}\boldsymbol{x} = \begin{bmatrix} m_0 m_1 m_2 \\ m_3 m_4 m_5 \\ m_6 m_7 m_8 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{11}$$

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 특징 점에 의하여 두 영상간의 대응관계를 계산할 수 있다. 따라서 검색된 중첩영역에서 4쌍의 대응 특징 점을 선택하여야 한다. 이 때 대응 특징 점은 중첩된 사각 영역의 8개의 꼭지 점이다.

(11)을 풀어쓰면 (12), (13)을 얻을 수 있다.

$$x' = \frac{m_0 x + m_1 y + m_2}{m_6 x + m_7 y + 1} \tag{12}$$

$$y' = \frac{m_3 x + m_4 y + m_5}{m_6 x + m_7 y + 1} \tag{13}$$

초기 변환행렬 M은 4쌍의 대응점을 통해 구할 수 있는 데 4쌍의 대응점에 대해 (12), (13)을 적용하여 정리하면 (14)를 얻을 수 있다.

$$\begin{bmatrix} x_1' \\ y_1' \\ x_2' \\ y_2' \\ x_3' \\ y_3' \\ x_4' \\ y_4' \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1'x_1 & -x_1'y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y_1'x_1 & -y_1'y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2'x_2 & -x_2'y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -y_2'x_2 & -y_2'y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3'x_3 & -x_3'y_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -y_3'x_3 & -y_3'y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4'x_4 & -x_4'y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -y_4'x_4 & -y_4'y_4 \end{bmatrix} \begin{bmatrix} m_0 \\ m_1 \\ m_2 \\ m_3 \\ m_4 \\ m_5 \\ m_6 \\ m_7 \end{bmatrix} \tag{14}$$

3.2 최적 변환 행렬의 계산

(14)를 간단한 수치해석법에 의해 풀면 변환 행렬 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-Marquradt 방법[1]을 이용하여 오차를 최소화한다. 이 방법은 중첩영역에서 최소가 되는 에러 e를 구해 새로운 변환 행렬 M을 결정하는 방법으로 미지의 운동 요소 $\{m_0,\cdots,m_7\}$ 에 대한 e의 편미분 값을 이용한다. 변환 전의 영상과 변환 후 영상간의 차이 e는 (15)와 같이 정의한다. 식에서 $I(x_i,\ y_i)$는 pixel의 명도 값(Intensity)을 나타낸다.

$$e = I'(x_i',\ y_i') - I(x_i,\ y_i) \tag{15}$$

최적 변환이란 (15)의 값이 최소가 되는 변환행렬 M을 구하는 것이다. 여기에서 8개의 대응되는 운동 요소의 미분은 (12)와 (13)로부터 다음과 같이 구할 수 있다.

$$\left.\begin{aligned}
\frac{\partial e}{\partial m_3} &= \frac{\partial e}{\partial m_0} = \frac{\partial x}{D}\frac{\partial I'}{\partial x'} \\
\frac{\partial e}{\partial m_4} &= \frac{\partial e}{\partial m_1} = \frac{\partial y}{D}\frac{\partial I'}{\partial x'} \\
\frac{\partial e}{\partial m_5} &= \frac{\partial e}{\partial m_2} = \frac{1}{D}\frac{\partial I'}{\partial x'} \\
\frac{\partial e}{\partial m_6} &= -\frac{x}{D}\left(x'\frac{\partial I'}{\partial x'} + y'\frac{\partial I'}{\partial y'}\right) \\
\frac{\partial e}{\partial m_7} &= -\frac{y}{D}\left(x'\frac{\partial I'}{\partial x'} + y'\frac{\partial I'}{\partial y'}\right)
\end{aligned}\right\} \tag{16}$$

(16)에서 D는 식(12)의 분모를 나타낸다.

위의 편미분으로부터 Levenberg-Marquardt 법을 이용하여 (17)과 같이 헷세 행렬 A와 가중 기울기 벡터(Weighted Gradient Vector) b를 계산할 수 있다.

$$A = a_{kl} = \sum_i \frac{\partial e_i}{\partial m_k}\frac{\partial e_i}{\partial m_l} \qquad b = b_k = -\sum_i \frac{\partial e_i}{\partial m_k} \tag{17}$$

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

이 결과를 이용하여 Levenberg-Marquardt 알고리즘의 변위 계산식 $\Delta m = (A+\lambda I)^{-1}b$ 에 대입하여 운동 변수Δm을 새로 계산하는 것이 가능하다. 여기에서 λ는 반복 계산에 대한 안정화 변수이다. 이 값은 (15)에서 오차를 측정했을 때 이전 단계에서 최소화를 수행한 것보다 에러가 증가하면 그 값을 감소시키고, 감소하면 증가시켜 새로운 운동변수 Δm을 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.

3.3 가중평균에 의한 블렌딩

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 픽셀 값으로 대신한다.

(18)과 (19)는 눈에 보이는 경계선을 줄이기 위한 가중평균의 x와 y방향에 대한 블렌딩 함수를 나타낸다.

$$w(x) = 1 - \left|\frac{2x}{width} - 1\right| \qquad w(y) = 1 - \left|\frac{2y}{height} - 1\right| \tag{18}$$

$$w(x,y) = w(x)w(y) \tag{19}$$

위 식을 이용하여 A, B 두 영상을 합할 경우 새로 합성된 영상(x,y)위치의 픽셀값은 (20)으로 결정된다.

$$I(x,y) = \frac{w_A(x,y)I_A(x,y) + w_B(x,y)I_B(x,y)}{w_A(x,y) + w_B(x,y)} \tag{20}$$

1) 이 방법은 비선형계의 최적치를 구하는 알고리즘이다.

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

$$I = \frac{\sum_k w_k(x_i, y_i) I_k(x_i, y_i)}{\sum_k w_k(x_i, y_i)} \tag{21}$$

결국 (21)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.

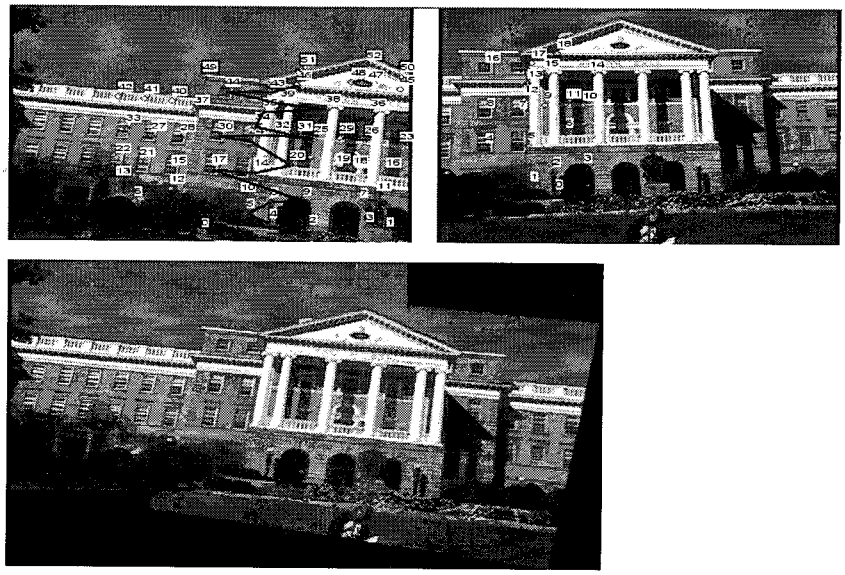


그림 10. 상하 이동, 회전 요소를 포함하는 실외 영상

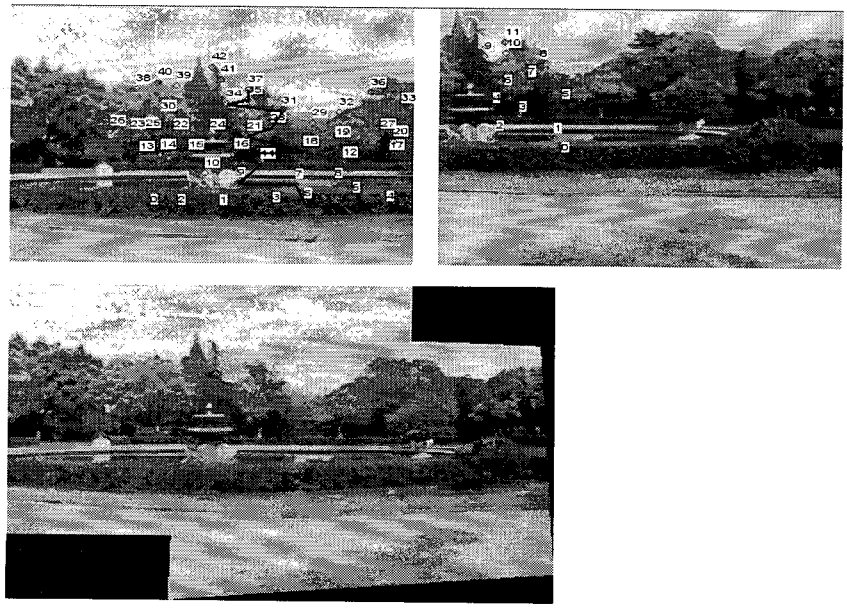


그림 11. 상하 이동 요소를 포함하는 실외 영상

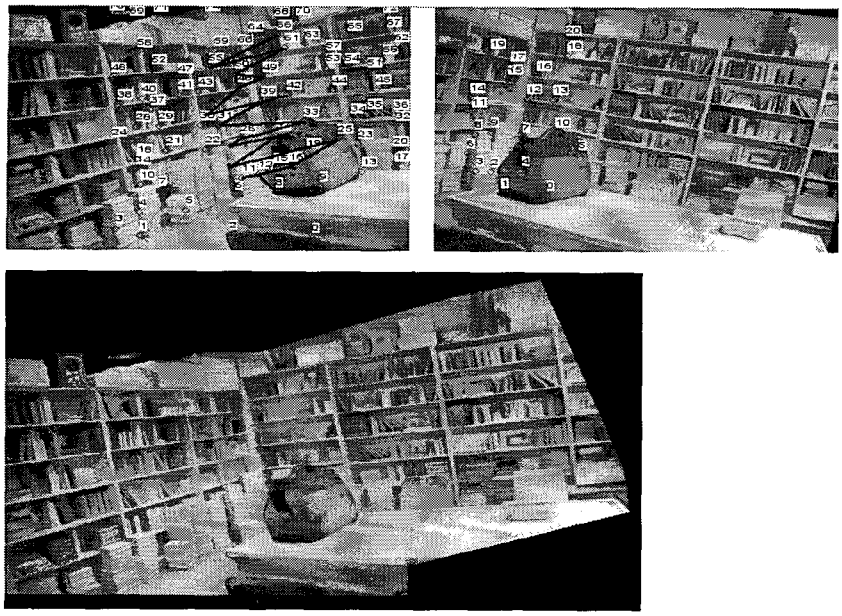


그림 12. 회전 요소가 있는 실내 영상

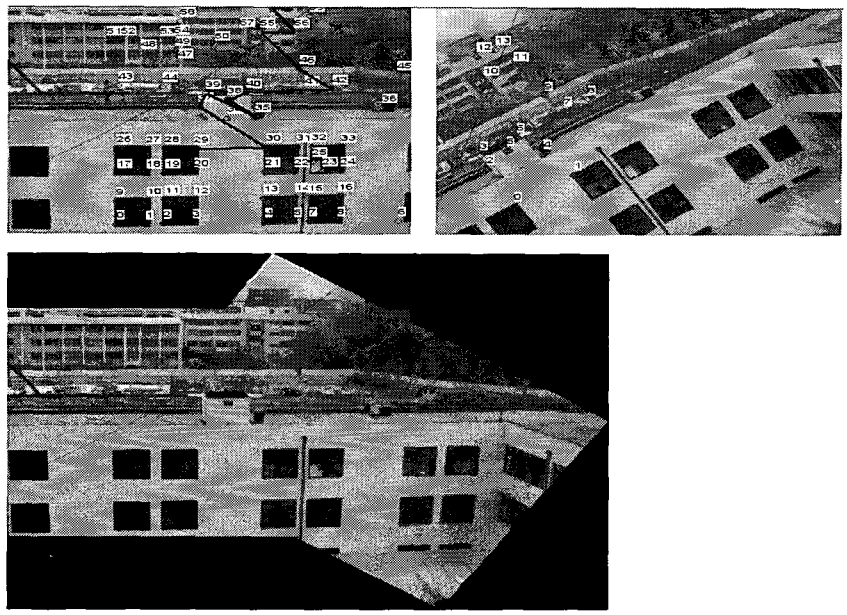


그림 13. 회전 요소가 있는 실외 영상

4. 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 회전 요소가 포함된 영상과 상하 이동요소가 포함된 영상, 회전요소와 상하이동요소가 포함되지 않은 실내, 실외 영상으로 나누어 실험하였다.

위의 실험 결과(그림 10~14)를 보면 5가지 경우

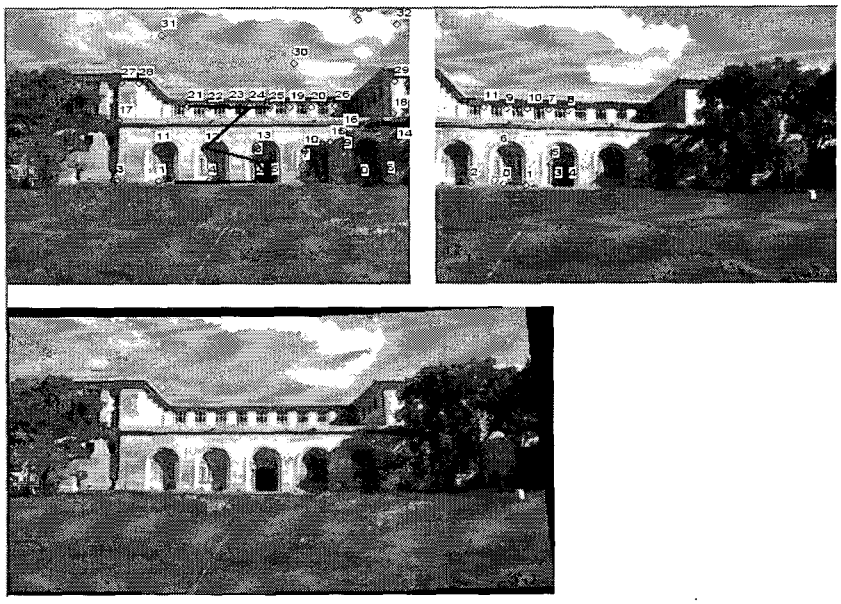


그림 14. 좌우 이동 요소만 있는 실외 영상

모두 안정된 결과를 보임을 알 수 있다. 특히 회전요소가 포함된 영상의 경우에도 정확한 대응점을 도출하여 정확한 모자익 영상을 생성하는 것을 볼 수 있다.

표 1은 본 논문에서 제안한 알고리즘의 우수성을 입증하기 위해 비교 항목에 해당하는 검색 시간에 대한 것을 검색 방법별로 자세히 나타낸 것이다. 실험 결과로 나타난 수치는 회전된 영상들과 좌우이동만 있는 영상들을 방법별로 알고리즘을 적용한 후 소요된 검색 시간을 측정한 것이다. 표 1의 결과를 보면 대부분의 방법들이 회전된 영상들의 경우에 좋지 않은 결과를 나타내고, 좌우 이동만 있는 영상들의 경우에도 검색시간이 많이 소요 되는걸 볼 수 있다.

그림 15는 수치 자료를 각 방법별로 검색시간을 구성하는 항목들을 합산하여 하나의 검색 시간으로 나타낸 것이다. 제안한 방법을 이용한 검색 방법이 검색 시간과 회전된 영상 모두 매우 좋은 결과를 보임을 알 수 있다.

본 논문에서는 비디오 카메라로 인접한 장면을 연결하여 촬영한 영상간의 중첩영역을 빠르고 정확하

표 1. 특징점 기반 영상 정합법의 검색 시간 비교(x표시는 검색 실패, 단위는 sec)

	영상 1	영상 2	영상 3	영상 4	영상 5	영상 6	영상 7	영상 8	영상 9	영상 10	영상 11	영상 12	영상 13	영상 14	영상 15
회전각도	8.8	14.9	25.6	30.9	49.6	-16.6	-22.1	-26.3	-31.7	-48.5	0	0	0	0	0
Intensity Matching	×	×	×	×	×	×	×	×	×	×	17.33	16.77	19.25	19.33	20.22
Hausdorff Matching	4.12	×	×	4.45	×	4.25	4.32	×	×	×	2.15	2.26	2.65	2.85	2.76
Statistics Matching	×	×	×	×	×	×	×	×	×	×	10.23	11.22	12.23	11.24	11.09
Optical-Flow Matching	4.76	3.98	5.94	×	×	3.63	3.34	5.93	×	×	1.89	2.36	3.62	2.78	2.79
보로노이 (이항검색)	2.35	×	×	×	×	×	×	×	×	×	1.56	2.27	1.78	2.34	1.34
보로노이 (기하위상기반)	0.075	0.078	0.076	0.078	0.067	0.063	0.075	0.076	0.075	0.063	0.055	0.067	0.056	0.055	0.066

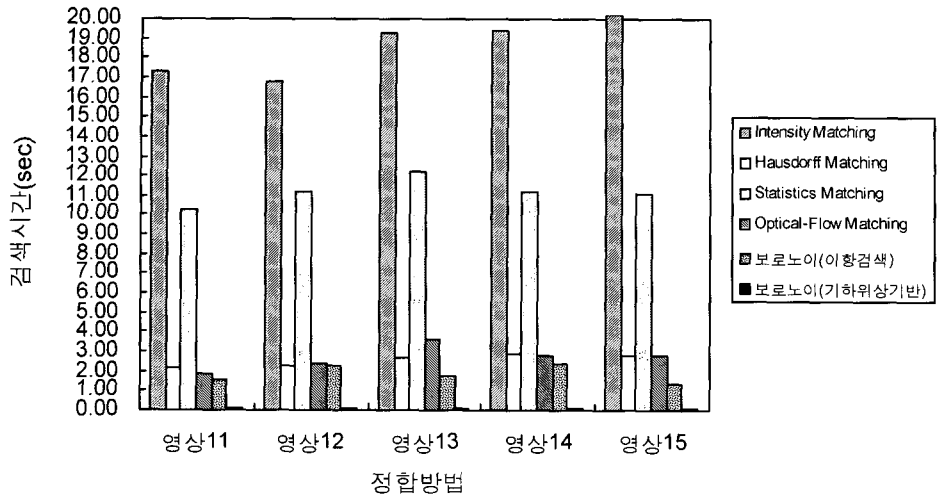


그림 15. 특징점 기반 영상 정합법의 검색 시간 비교 그래프

게 검색하는 방법에 대하여 기술하였고, 영상내의 특징 점들의 기하 위상 정보와 특징점 사이의 거리 정보를 가지는 보로노이 거리(Distance Transform)의 선형적인 특성에 기반한 형태 정보 매칭 기법을 사용하였다. 영역간의 구조적인 유사도와 형태정보를 이용함으로써 빠르면서 회전 불변하는 중첩영역 검색이 가능하도록 하였다.

향후 연구 방향은 스케일이 다른 영상과 변형된 영상의 중첩영역 검색이 가능케 하는 방법과 시간을 효과적으로 줄이는 방법에 대하여 연구할 계획이다.

참 고 문 헌

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

[2] Daniel P. Huttenlocher and 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] Naoki Chiba, Hiroshi Kano, Minoru Higashihara, Masashi Yasuda, and Osumi, "Feature-Based Image Mosaicing," *MVA'98 IAPR Workshop on Machine Vision Applications*, pp. 5-10, 1998.

[4] 고종호, 이칠우, "특징기반 영상 모자이킹," *IPIU2000(제 12회 영상 처리 및 이해에 관한 워크샵)*, pp. 275-280, 2000. 1.

[5] G. Borgefors, "Hierarchical Chamfer Matching; A Parametric Edge Matching Algorithm," *IEEE Trans, Pattern Analysis and Machine Intelligence*, Vol. 10, No. 6, pp. 849-865, Nov. 1988.

[6] Zhang Zhijia, Huang Shabai, and Shi Zelin, "A fast strategy for image matching using Hausdorff distance" *Robotics, Intelligent Systems and Signal Processing, 2003. Proceedings. 2003 IEEE International Conference*, Vol. 2 pp. 915-919 8-13 Oct. 2003.

[7] Dan Huttenlocher and Phil Torr, "Efficient Algoritms for Matching" *Int'l Conf ICCV* 2003.

[8] Axel Pinz, Mafred Prantl, and 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.

[9] Chris Harris and Mike Stephens. A Combined Corner and Edge Detector. *In Proc. 4th Alvey Vision Conference*, Manchester, pages 147-

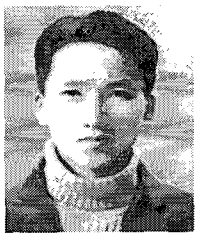
151, 1988.

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

[11] 정민영, 이칠우, 배기태, 이동휘, "보르노이 거리 정합을 이용한 영상 모자익" *멀티미디어 학회 논문지*, 제6권 제7호, pp. 1179-1188 2003. 12.

[12] Guest, E, Berry, E, Baldock, R.A, Fidrich, M, and Smith, M.A, "Robust point correspondence applied to two- and three-dimensional image registration" *Pattern Analysis and Machine Intelligence, IEEE Transactions on* Vol. 23, Issue 2, pp. 165-179, Feb. 2001.

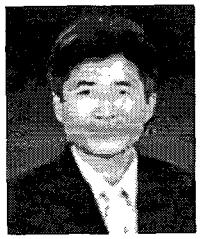
[13] Triggs, B, "Joint feature distributions for image correspondence" *Computer Vision, ICCV 2001. Proceedings. Eighth IEEE International Conference*, Vol. 2 pp. 201-208, 7-14 July 2001.



배 기 태

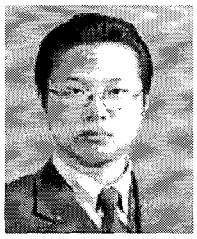
1997년 호원대학교 전자계산학과 졸업(학사)
1999년 전남대학교 대학원 컴퓨터공학과(공학 석사)
2001년~현재 전남대학교 대학원 컴퓨터공학과(박사 수료)

관심분야 : 컴퓨터 비전, 멀티미디어 데이터베이스, 패턴 인식, 컴퓨터 그래픽스, 영상 처리



정 민 영

1991년 숭실대학교 전자계산학과 졸업(학사)
1993년 숭실대학교 대학원 전자계산학과(공학 석사)
2004년 전남대학교 대학원 컴퓨터공학과(공학 박사)
1993년~1996년 숭실대학교 중앙전자계산소 연구원
1997년~1999년 광주여자대학교 정보전산원장
1996년~현재 광주여자대학교 멀티미디어학과 교수
관심분야 : 컴퓨터 비전, 멀티미디어 데이터베이스, 컴퓨터그래픽스, 소프트웨어 공학



오 재 용

2000년 전남대학교 컴퓨터공학과 졸업(학사)
2002년 전남대학교 대학원 컴퓨터공학과(공학 석사)
2002년~현재 전남대학교 대학원 컴퓨터공학과 박사과정

관심분야 : 컴퓨터 비전, 제스처 인식, 컴퓨터 그래픽스

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
1996년~현재 전남대학교 전자컴퓨터정보통신 공학부 교수
관심분야 : 컴퓨터 비전, 멀티미디어 데이터베이스, 컴퓨터그래픽스