

인지기반 영상처리를 위한 특징점 검출 및 응용 기술

원치선·엄기현·조경은 (동국대학교), 심성대 (국방과학연구소)

I. 서론

시각적 특이성(visual saliency)이란 영상내에 존재하는 객체들 사이의 상대적인 특이성으로 <그림 1>-(a)의 흑점은 다른 점들과 비교해서 상대적으로 높은 특이성을 나타내는 반면 <그림 1>-(b) 내의 흑점은 (a)의 그것보다 상대적인 특이성이 떨어진다. <그림 1>-(c)에서는 특정 점의 특이성보다 점들 사이의 상호 동등성이나 화합성이 더 강하다^[1]. 영상 내의 특이성을 표현할 수 있는 물리적인 실체는 <그림 1>과 같이 각 픽셀이 갖는 칼라의 분포와 연결성(에지) 혹은 같은 밝기나 패턴을 갖는 픽셀 군집의 모양 등 다양하다. 이들 여러 가지의 특징점을 통합하여 영상 내에서 상대적인 특이성을 시각적으로 표기하여 특징점 지도(Salient Map)로 나타낼 수 있다.

특징점 지도를 생성하기 위해 영상 내에 존재하는 국부적인 특징(local feature) 정보를 추출할 필요가 있다. 국부적인 특징은 특정 픽셀이나 픽셀 그룹의 칼라, 에지 혹은 모양이 이웃하는 다른 픽셀들의 그것과 의미 있는 차이를 나타낸다. 일단 특징점들이 추출되면 특징점을 중심으로 주변 픽셀의 특징 정보를 수치화하여 해당 특징점의 서술자(descriptor)를 생성할 수 있으며 이를 바탕으로 영상 검색, 영상 등록(registration), 객체 추적, 그리고 객체 인식 등 다양한 컴퓨터 비전 문제에 활용될 수 있다. 그러나 카메라 정렬 등과 같은 응용분야에서는 특징점 주변의 픽셀들의 정보(특징점의

서술자)는 무시되고 해당 특징점의 위치 정보만 필요한 경우도 있다. 어떤 경우든지 좋은 특징점은 다음과 같은 기준을 만족하여야 한다^[2].

- 재연성 (repeatability)

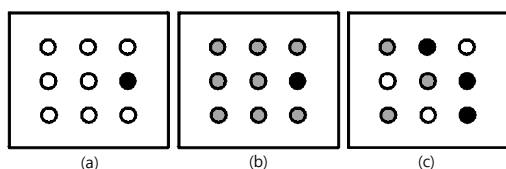
동일한 물체를 포함하지만 영상의 취득 조건(혹은 시점)이 상이한 두 영상에 대해 재연성이 높을수록 특징점이 양쪽 영상에서 동시에 검출되는 비율이 높다.

- 유익성/차별성 (informativeness/distinctiveness)

검출된 특징점들이 특징벡터의 공간상에 광범위하게 퍼지는 기술은 특징점들 사이에 정보의 중복성이 낮고 각 특징점이 전달하는 정보의 유익성이 높다.

대부분의 응용에서 재연성은 중요한 요구조건으로 높은 재연성을 갖는 검출기는 불변성(invariance)과 강인성(robustness)을 필요로 한다. 즉, 영상에 가해질 수 있는 다양한 변형에 대해 검출의 결과가 영향을 받지 않도록 하는 불변성과 잡음이나 압축 부작용 등에 대한 민감도를 낮추는 강인성이 필요하다. 그 밖에 검출된 특징점의 개수도 좋은 검출기의 기준에 포함될 수 있다. 응용에 따라서는 특정 객체에 대해 충분히 많은 특징점이 검출되어야 한다. 하지만 또 다른 응용에서는 불필요하게 많은 특징점이 계산 복잡도를 늘리는 등 부정적으로 작용할 수 있다. 마지막으로 검출된 특징점의 정확도(위치)와 수행시간(계산 복잡도)도 좋은 검출기를 판정하는 기준이 된다.

본 기고에서는 영상내의 국부적인 특징점을 추출하는 핵심적인 기술과 응용 사례에 대해 소개한다. 영상처리 및 컴퓨터 비전을 위한 특징점(salient point)은 관심점(interesting point), 코너(corner), 그리고 블럽(blob) 등을 포함한다. 이들



<그림 1> 시각적 특이성의 예



은 모두 대표적인 국부 특징점들이며 인간 시각 심리학적으로 객체의 인지에 큰 영향을 미친다. 예를 들어, 어떤 객체의 코너 점들을 연결하는 경계선의 상당 부분이 실종되어도 경계선들이 모여서 생성된 코너와 그 주변의 정보만 주어지면 해당 객체의 모양 정보를 어느 정도 인지할 수 있다. 따라서 특징점에서 코너가 기여하는 비중이 크며 본 기고에서는 주로 코너를 특징점으로 추출하는 방법에 집중한다. 코너는 영상객체의 모서리에 존재하는 특징점이지만 추출된 코너는 반드시 실제의 코너일 필요는 없으며 주변 픽셀과 구별되는 특징(예를 들어, 동일 픽셀에 두 개의 서로 다른 방향성 에지가 존재하는 경우)으로 정의될 수 있다. 하지만 많은 경우 코너를 광의의 특징점으로 다루며 위에서 언급한 재연성과 유익성을 갖춘 곳이라면 그것이 실제 코너이든 아니든 상관없이 특징점으로 인정받을 수 있다.

II. 기본적인 검출기

1. Harris 코너 검출기

에지의 검출에 기반을 둔 특징점은 기본적으로 국부적인 밝기의 구배(gradient)를 이용한다. 즉, 각 픽셀을 중심으로 일정 크기의 창(window)을 설정하며 창 내의 모든 픽셀들의 구배 정보로부터 해당 픽셀을 중심으로 복수개의 에지가 만나면 코너로 간주한다. 예를 들어, Harris 검출기^[3]는 각 픽셀에 대해 서로 다른 방향의 에지가 존재하는지를 판단하여 서로 다른 방향성의 에지가 한 점에서 만나는 경우 코너로 정의한다. 이를 수식적으로 나타내면 픽셀 t 위에 Gaussian 가중치를 갖는 창 W 를 위치한 후 창 내의 모든 픽셀 (x,y) 에 대해 $(\Delta x, \Delta y)$ 의 위치 이동에 따른 아래의 식과 같은 밝기 변화의 자기상관함수 (autocorrelation function) C_t 를 기반으로 코너의 존재 여부를 결정한다.

$$C_t = \sum_{(x,y) \in W} (I(x,y) - I(x + \Delta x, y + \Delta y))^2 \quad (1)$$

식 (1)에서 $I(x,y)$ 는 (x,y) 에서의 밝기값을 나타내며 $\sum_{(x,y) \in W}$ 는 창 W 내에 Gaussian 가중치를 고려한 합을 표시한다. 식(1)은 특정 방향 $(\Delta x, \Delta y)$ 에 대한 구배의 강도를 나타내지만 좌표축 변환을 통해 픽셀 t 주변에서 밝기의 변화가 가장 큰 축과 가장 작은 축으로 나타낼 수 있으며 이를 위해 식 (1)의 $I(x+\Delta x, y+\Delta y)$ 를 Taylor 전개식으로 근사화하면 아래와 같이 매트릭스 방정식으로 나타낼 수 있다.

$$C_t = [\Delta x \ \Delta y] \begin{bmatrix} \sum_W (I_x(x,y))^2 & \sum_W I_x(x,y)I_y(x,y) \\ \sum_W I_x(x,y)I_y(x,y) & \sum_W (I_y(x,y))^2 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (2) \\ = [\Delta x \ \Delta y] M_t \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

식 (2)에서 $I_x(x,y) = \partial I(x,y) / \partial x$, $I_y(x,y) = \partial I(x,y) / \partial y$ 이며 2×2 매트릭스 M_t 는 픽셀 t 주변 밝기의 구배에 대한 정보를 포함한다. M_t 가 두 개의 서로 다른 고유치(eigenvalue)를 갖는다면 M_t 는 고유치를 대각 원소로 하는 대각선 매트릭스(diagonal matrix) D 로 대각화되며 $M_t = PDP^{-1}$ 로 나타낼 수 있다. 이 표현을 식 (2)에 대입하면 일정한 C_t 값에 대해 식(2)는 타원형을 나타내며 타원의 장반경(semi-major axis)과 단반경(semi-minor axis)은 각각 밝기의 변화가 가장 빠른 방향과 가장 느린 방향을 나타내게 된다. 따라서 타원 반경에 직접적인 영향을 주는 M_t 의 두 고유치의 절대적 및 상대적인 값에 따라 평탄, 에지 및 코너 점으로 분류될 수 있다. 즉, 두 고유치가 모두 일정 크기 이상이면 타원의 장반경과 단반경에서 모두 빠르게 밝기가 변하고 곡률이 커지게 되며 이 경우 픽셀 t 는 서로 다른 두 방향의 에지가 만나는 점으로 볼 수 있는 코너의 특징점으로 인정된다. 고유치를 계산하는 과정을 생략하여 코너 응답값 $R = |M_t| - k(\text{Trace}(M_t))^2$ ($k=0.04 \sim 0.06$)을 사용할 수 있으며 특정 픽셀의 응답값 R 이 주변의 그것들보다 큰 경우 그 픽셀을 코너로 결정할 수 있다^[3].

2. SUSAN 검출기

Harris 코너 검출기는 기본적으로 밝기의 구배 정보(미분기)를 이용하므로 잡음에 민감하다. 이와 같은 문제점을 극복하기 위해 각 픽셀의 반경 내에 유사 밝기를 갖는 픽셀과 그렇지 않은 픽셀들의 면적의 비율로 코너를 검출하는 SUSAN (Smallest Univalve Segment Assimilating Nucleus)^[4] 방법이 제안되었다. 즉, 특정 픽셀을 중심으로 일정 반경 내의 픽셀들이 대부분 서로 유사하면 해당 픽셀은 평탄영역에 속하고 두 개의 서로 다른 밝기의 값이 대략 절반 정도 분포하면 에지(edge)로 볼 수 있고 어느 한 쪽이 다른 밝기 값의 분포에 비해 월등히 많으면 해당 픽셀을 코너로 판정한다.

3. Harris-Laplace 및 Harris-Affine 검출기

Harris 코너 검출기와 SUSAN 검출기는 수평이동과 회전 등에 대해 강인하고 재연성이 높으나 스케일(scale) 및 아핀(affine) 변환에 강인하지 못하다. 이와 같은 문제점을 해결하기 위해 스케일 변환에 불변인 Harris-Laplace 검출기와 아핀 변환에 영향을 받지 않는 Harris-Affine 검출기가 제안되

었다^[5].

스케일 변환에 대해 강인한 대부분의 특징점 검출기는 스케일-공간 양쪽 도메인에서 동시에 특징점으로 검출된 점들만을 최종 특징점으로 인정하는 방법을 기반으로 한다. 이를 위해 우선 Harris 코너 검출기와 같이 기존의 (스케일 변환에 약한) 검출기로 특징점을 검출하고 그 중에 주어진 *스케일 선택함수*가 특정 스케일에서 특히 강한 특징성(즉, 극값)을 나타내는 특징점만 속아내는 과정으로 이루어진다. 예를 들어, Harris 코너 검출기로 검출된 특징점들에 대해 다양한 스케일에서 스케일 선택함수를 적용한 결과 <그림 2>-(a)에서와 같이 특정 스케일 σ_1^* 에서 극값(extreme value)을 갖는다면 해당 특징점은 영상의 크기가 줄어든 경우에도 같은 점에서 <그림 2>-(b)와 같이 또 다른 스케일 σ_2^* 에서 극값을 갖게 될 가능성이 크며 따라서 스케일 변환에 영향을 받지 않고 항상 특징점으로 검출될 수 있다.

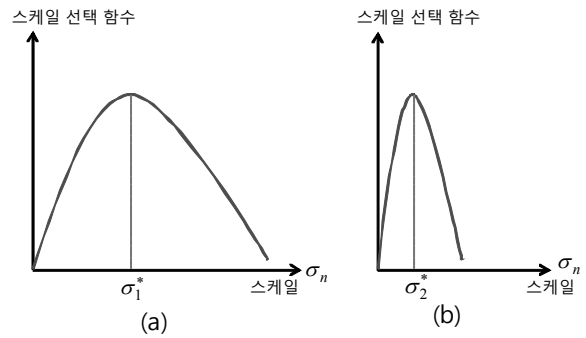
Harris-Laplace 검출기는 Harris 검출기로 초기 특징점을 찾고 각 특징점에 대해 Laplacian 2차 미분 함수(Laplacian 응답)를 스케일 선택함수로 사용하여 최종 특징점을 속아내는 방법이다. 즉, 식 (2)의 2x2 Harris 매트릭스 M_t 를 사용하여 특징점 후보를 검출한 후 각 특징점 후보들에 대해 Gaussian 분산값(σ_n)을 증가시키면서 (즉, Gaussian 창 W 확대) 여러 스케일에서 아래의 LoG (Laplacian of Gaussian) 필터($LoG(\sigma_n)$)와 특징점 후보의 주변 영상 데이터 세 대해서로 컨벌루션(convolution)을 취한 후 그 절대치가 극값을 갖는 스케일(σ_k^*)이 존재하면 해당 코너는 스케일 변화에 영향을 받지 않는 특징점으로 간주된다.

$$LoG(\sigma_n) = \sigma_n^2 (G_{xx}(x, y, \sigma_n) + G_{yy}(x, y, \sigma_n)) \quad (3)$$

식 (3)에서 $G_{xx}(x, y, \sigma_n)$ 와 $G_{yy}(x, y, \sigma_n)$ 는 각각 2차원 Gaussian 함수($G(x, y, \sigma) = \exp\{-(x^2 + y^2)/2\sigma^2\}/(2\pi\sigma^2)$)에 대한 x축 및 y축의 2차 미분을 의미한다¹⁾. 식 (3)의 LoG는 원형 대칭성을 가지며 따라서 블럽(blob) 형태의 국부적인 구조를 포함하는 스케일에서 그 절대값이 최대가 된다. 따라서 만약 영상의 크기에 변화가 발생하면 해당 블럽의 크기에도 변화가 발생하지만 크기가 바뀐 블럽은 또 다른 스케일에서 극값을 가지며 검출될 가능성이 높으므로 스케일의 변화에 영향을 받지 않게 된다.

영상에 아핀(affine) 변환이 가해지면 방향 별로 스케일 변

1) 에지에 대한 Gaussian 필터의 미분값은 σ 에 반비례하며 따라서 스케일 변환을 위해 σ 를 증가시키면 Laplacian 응답은 감소한다. 이를 보상하기 위해 식 (3)에 σ_n^2 가 곱해졌다. 이와 같이 σ_n^2 가 곱해진 Laplacian을 Normalized Laplacian이라고도 한다.



<그림 2> 스케일 선택함수가 극값을 갖는 예 (a) 영상의 크기가 줄어들기 전, (b) 영상의 크기가 줄어든 후

환의 크기가 고르지 않고 서로 달라질 수 있다. 따라서 모든 방향에 동일한 스케일 변환을 위해 고안된 Harris-Laplace 검출기는 적용될 수 없다. 즉, 각 스케일마다 유지되어 오던 원형의 창을 타원형을 포함하도록 일반화되어야 한다. 이 경우 원형 창에서 계산된 매트릭스 M_t 는 가해진 아핀 변환에 맞게 타원형 내의 매트릭스로 변형되어야 하며 이를 위해 Harris-Laplace 검출기로 검출된 초기 특징점을 반복적으로 타원형의 아핀 변환에 해당되는 매트릭스로 변환 후 특징점을 보정해 나가는 Harris-Affine 검출기가 제안되었다^[5].

4. Hessian 기반 검출기

식 (3)의 LoG는 원형 대칭성을 가지므로 (blob) 형태의 국부적인 구조를 검출하는데 유용하다. 따라서 LoG을 일반화한 형태로 스케일 σ_n 의 픽셀 t에서의 Hessian 매트릭스 $H_t(\sigma_n)$ 도 식 (2)의 Harris 매트릭스 M_t 와 같이 특징점을 검출하는 기본 매트릭스로 활용될 수 있다.

$$H_t(\sigma_n) = \begin{bmatrix} \sum_W I(x, y) G_{xx}(x, y, \sigma_n) & \sum_W I(x, y) G_{xy}(x, y, \sigma_n) \\ \sum_W I(x, y) G_{xy}(x, y, \sigma_n) & \sum_W I(x, y) G_{yy}(x, y, \sigma_n) \end{bmatrix} \quad (4)$$

식 (4)에서 \sum_W 는 현재 픽셀을 중심으로 창 W내의 모든 픽셀 (x, y)에 대한 영상데이터 $I(x, y)$ 와 Gaussian의 2차 미분의 곱을 합산한 것이며 W내의 합에 Gaussian 가중치를 고려하는 식 (2)의 그것과는 구별된다. $H_t(\sigma_n)$ 의 행렬식(determinant)과 Trace는 영상의 국부적인 구조를 추출하는 필터로 활용될 수 있다. 예를 들어, $H_t(\sigma_n)$ 의 Trace는 LoG 필터의 특성을 가지며 블럽 형태의 국부적인 구조를 검출 할 수 있다. 하지만 LoG (혹은 DoG) 만을 적용했을 때 특징점 검출을 나타내는 극값이 곡선이나 에지의 점에서도 발생하는 단점이 있다. 따라서 특징점의 검출 성능을 향상시키기 위해 $H_t(\sigma_n)$ 의 Trace 뿐만이 아니라 행렬식에 의해서도 모두 극점을 나타

내는 픽셀을 특징점으로 인정하도록 기준을 강화할 필요가 있다.

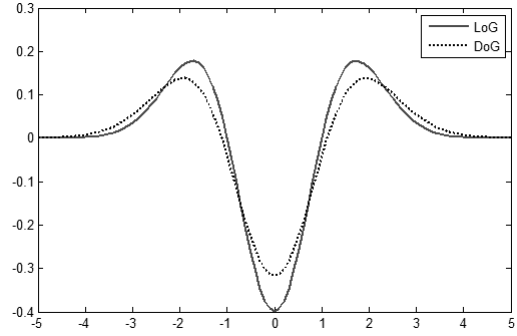
즉, Hessian 매트릭스 $H(\sigma_n)$ 의 행렬식과 Trace가 모두 극값을 갖는 픽셀을 특징점으로 인정하는 기준을 채택하여 새로운 특징점 검출기로 사용할 수 있으며 Harris 매트릭스 M_t 를 대신할 수 있다. 따라서 스케일과 아핀 변환에 강인하도록 조합된 Harris-Laplace 및 Harris-Affine 검출기 대신 Hessian-Laplace 및 Hessian-Affine 검출기를 고려할 수 있다^[5]. 그러나 Harris 매트릭스를 사용하는 경우 코너를 중심으로 특징점을 검출하는 경향이 있는 반면에 Hessian 매트릭스를 사용하는 경우 블립을 중심으로 특징점을 검출하는 경향이 강하다.

Hessian 매트릭스를 스케일 불변 검출기로 활용하기 위해 다양한 스케일 (σ_n)에 적용할 필요가 있다. 그러나 σ_n 을 증가시키면 식 (4)의 필터창의 크기도 증가하여 많은 계산을 필요로 한다. 어차피 디지털 영상에 대한 이산(discrete) 연산을 위해 식 (4)의 2차 미분 Gaussian 필터의 계수도 이산화 및 양자화가 필요하며 이 과정에서 에러가 발생하고 정확한 필터의 구현에 한계가 있다. 따라서 Hessian 매트릭스 $H(\sigma_n)$ 의 2차 Gaussian 미분 필터 계수를 근사화하여 박스필터(box filter)를 적용하고 적분영상(integral image) 개념을 도입하여 창 크기에 상관없이 계산량을 대폭 줄인 SURF(Speed-up robust features)^[6] 검출기가 제안되었다.²⁾

5. Scale Invariant Feature Transform (SIFT) 검출기

SIFT 검출기^[7]는 Harris-Laplace 및 Harris-Affine과 같이 공간 영역에서 선택된 특징점 후보들 중에 다양한 스케일에서 스케일 변환에 영향을 받지 않을 특징점만을 선택한다는 점에서 서로 유사하다. Harris-Affine과 달리 SIFT는 아핀 변환에 완벽하게 대응하지는 못하지만 부분적으로 아핀 변환에 영향을 받지 않는다. 반면에 많은 특징점을 신속하게 검출하는 장점을 가진다. 이와 같은 검출기 효율성(수행속도의 향상)은 식 (3)에서 사용한 스케일 선택함수 LoG를 더 간단한 DoG (Difference of Gaussian)로 근사화하여 대체 적용함으로써 달성된다.³⁾

- 2) 적분영상은 영상좌표의 원점으로부터 영상데이터를 포함하는 사각형을 확장하며 영상데이터 값을 누적한 것으로 Hessian 검출기에 사용하는 박스필터(필터계수를 근사화하여 일정 필터영역(박스) 내에 동일한 계수값을 사용한 필터)의 크기에 상관없이 빠른 시간에 필터링을 수행할 수 있다.
- 3) 좀 더 구체적으로 1차원 신호에 대해 $DoG(x, k, \sigma) \approx (k-1)LoG(\sigma)$ 의 관계로부터 <그림 3>에 $\sigma = 1, k = 2^{1/3} \approx 1.26$ 의 경우 DoG와 LoG의 유사성을 확인할 수 있다.



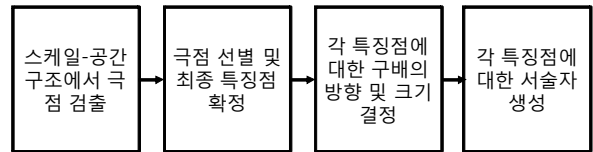
<그림 3> LoG와 DoG의 비교 ($\sigma = 1, k = 2^{1/3} \approx 1.26$ 을 갖는 1차 Gaussian의 경우)

$$DoG(x, y, k, \sigma) = G(x, y, k\sigma) - G(x, y, \sigma) \quad (5)$$

주어진 영상에 대해 식 (5)와 같이 Gaussian 필터의 σ 를 k배하고 차분 영상

$$D(x, y, k, \sigma) = DoG(x, y, k, \sigma) * I(x, y) = G(x, y, k\sigma) * I(x, y) - G(x, y, \sigma) * I(x, y) \quad (6)$$

를 얻고 이후 $k\sigma = 2\sigma$ 가 될 때까지 연속적인 차분영상들을 생성한 후 k번째 차분 영상의 전(k-1) 및 후(k+1)의 각 픽셀을 중심으로 3x3x3의 입체적인 이웃 픽셀들의 값들 중에 현재(k번째 차분 영상의 중심) 픽셀의 값이 3x3x3내의 26개의 모든 다른 픽셀의 값들보다 크거나(최대값) 혹은 작으면(최소값) 현재 픽셀을 특징점의 후보로 선정한다. 이후 선정된 특징점 후보 픽셀들 중에 밝기의 대비가 낮거나⁴⁾ 에지에 걸쳐있는 후보⁵⁾들을 제거하면 최종 특징점이 정해진다. 최종적으로 특징점들이 결정되면 각 특징점들의 주변 픽셀값들의 1차 미분의 크기와 방향성 정보를 이용하여 각 특징점들에 대한 서술자(descriptor)⁶⁾ 벡터를 생성하여 특징점들 사이의



<그림 4> SIFT의 전체 블록도

- 4) 특징점의 각 후보에 대해 정수 픽셀들 사이의 sub-pixel의 단위로 극점 값의 정확한 위치를 찾은 후 해당 sub-pixel에서의 D가 일정 값 이하인 특징점 후보들을 제거한다.
- 5) 각 특징점 후보들을 중심으로 2x2 Hessian 매트릭스를 구하고 Hessian 매트릭스의 두 고유치의 비가 주어진 값을 초과하면 좌표축 변환에 의해 생성된 타원의 장반경(semi-major axis)과 단반경(semi-minor axis)의 반경의 비율이 크고 따라서 해당 특징점은 에지에 걸쳐있는 것으로 판단한다.
- 6) 예를 들어 주어진 특징점에 해당되는 스케일의 Gaussian 필터 적용된 영상에 대해 해당 특징점 주변 16x16 픽셀을 4x4 블록으로 나누고 각 블록의 구배의 방향을 8개로 양자화하여 각 블록에 8개 bin의 히스토그램을 얻으면 128개의 원소를 갖는 특징벡터를 생성할 수 있다^[7].

〈표 1〉 대표적인 특징점 검출기 요약

기술명	검출원리	특성	기타
Harris 검출기	- 각 픽셀 주변의 윈도우내 미분값을 기반으로 복수의 에지가 존재하는 경우 코너로 정의	- 수평 이동 및 회전에 영향을 받지 않음 - 다양한 조명 조건에 강함	- 기본적으로 미분기를 사용하므로 잡음에 민감함
SUSAN 검출기	- 각 픽셀의 반경 내에 유사 밝기를 갖는 픽셀과 그렇지 않은 픽셀들의 분포의 비율로 코너 검출	- 수평 이동 및 회전에 영향을 받지 않음	- 미분기를 사용하지 않으므로 잡음에 강함
Harris-Laplace 검출기	- 다양한 스케일의 영상 변환에 Harris 검출기를 적용하고 그 중에 Laplacian 함수가 극값을 갖는 특징점을 코너로 인정	- 영상에 대한 수평 이동 및 회전 뿐만 아니라 스케일 변화에도 강인함	- 영상의 스케일 변환이 모든 방향에 동일하게 가해진 경우에 사용 가능
Harris-Affine	- Harris-Laplace 검출기로 검출된 초기 특징점의 위치를 가해진 아핀 변환에 맞게 반복하여 보정함	- 영상에 대한 수평 이동 및 회전뿐만 아니라 아핀 변환에도 강인함	
SIFT	- Harris-Laplace 변화과 같이 스케일-공간 영역에서 극값을 갖는 특징점을 찾음 - 스케일 선택함수로 LoG 대신 DoG 사용함	- 회전 및 스케일 변화에 영향을 받지 않음. 부분적인 아핀 변환에 강인	- 검출된 특징점 주변의 픽셀들의 구배의 방향과 크기로 특징점 매칭을 위한 서술벡터를 생성 - 여러 응용에 가장 빈번하게 활용되는 검출기임
Hessian 기반 검출기	- Laplacian 필터의 일반형으로 Gaussian 함수의 2차 미분 계수로 이루어진 2x2 Hessian 매트릭스의 행렬식과 Trace에 모두 극값을 갖는 픽셀을 특징점으로 함.	- LoG나 DoG에서 에지나 꼭선의 가운데 점을 특징점으로 찾는 단점 해소	- Hessian 필터 계수를 근사화하여 계산량을 대폭 줄인 효율적인 SURF 사용

유사도 판정에 활용된다. SIFT의 전체적인 블록도는 〈그림 4〉와 같다. 11장에서 살펴본 대표적인 특징점 검출기를 〈표 1〉에 요약하였다.

III. 특징점 검출기의 응용

1. 영상등록 (Image Registration)

영상등록은 복수의 영상으로부터 얻은 특징점들을 공통의 좌표에 사상하는 과정으로 겹친 영상들의 픽셀을 서로 정렬하거나 다양한 영상 소스 (예, 칼라, 레이저, 적외선, 양안 카메라 등)로부터 얻은 영상들을 동일 좌표계에 정렬하는 것이다. 따라서 영상 등록은 컴퓨터비전의 여러 문제를 해결하는 핵심적인 전처리 단계이다.

영상등록은 특징점 추출, 특징점 정합, 변환식 모델링, 그리고 영상 재표본(resampling) 과정으로 이루어져 있다. 이들 과정 중에 특징점의 추출과 정합은 변환 모델을 결정하는데 중요한 역할을 한다. 만약 정합 대상 영상들이 같은 시간에 같은 센서에서 취득되었다면 변환식을 정확히 얻어질 수 있지만 영상들 사이에 시간차이가 있거나 서로 다른 영상센서로부터 얻은 경우에는 SIFT 같은 기존의 특징점을 쓰더라도 정합에 오류가 발생할 수 있다.

레이저 센서를 이용한 경우 주변의 물체를 모두 포함하는 전체 피사체를 한 번의 레이저 스캔으로 취득할 수는 없고 보통 여러 번의 스캔을 통해 데이터 취득을 수행한다. 따라서 각각의 스캔에서 얻은 점군(point cloud) 데이터를 공통의 좌

표에 등록하는 과정이 필요하다. 즉, 스캔마다 부분적으로 겹치는 점군의 데이터를 사이에 변환 파라미터를 예측하여 보정할 수 있다. 그러나 최근에 대부분의 레이저 스캐너들이 해상도가 더 높은 칼라 영상 카메라를 같이 장착하고 있으므로 스캔된 영상들을 등록하기 위해 칼라 영상 데이터를 활용할 수 있다. 즉, 사실적인 3D 지도를 생성하기 위해 레이저 스캔으로 얻은 3D 점군 데이터를 2D 컬러 영상에 투영할 필요가 있으며 이를 위해 2D 영상의 SIFT 특징점의 서술자를 3D 점군과 정합하여 등록하는 기술이 필요하다. 예를 들어, [8]에서 SIFT 특징점 검출기를 이용하여 칼라 영상 기반 레이저 스캔 등록 방법이 제안되었다.

2. 영상 브라우징

인터넷에서 용이하게 얻을 수 있는 특정 객체(유명한 관광지 와 같은)에서 촬영된 수많은 영상 데이터를 바탕으로 사진 촬영시 사진사의 위치와 방향 정보를 파악하고 영상들을 서로 끊임없이 연결해줌으로써 해당 객체를 3D 환경에서 브라우징할 수 있는 환경을 제공한다. 이를 위해 각 영상들 사이에 정렬이 필요하며 SIFT 특징점 검출기를 적용하였다. SIFT는 영상 변환에 강인할 뿐만 아니라 각 특징점에서 서술자가 제공되어 영상들 사이에 특징점의 정합이 용이하다^[9].

3. 영상 재촬영

영상 재촬영 (Rephotographing)은 동일한 시점에서 기존

의 오래된 사진과 동일한 사진을 얻는 기술로서 역사적인 유물(건축물)의 사진과 동일한 시점에서 찍은 현재의 사진을 필요로 할 때 사용하는 기술이다. 기존의 역사적인 사진의 촬영 시점을 파악하기 위해 촬영시점으로 추측되는 위치에서 찍은 영상과 그곳으로부터 약 20도 회전한 시점에서 찍은 사진으로부터 3D 구조를 계산하고 촬영시점으로 추측되는 시점에서 찍은 사진과 기존의 사진과 특징점 정합으로 정렬한 후 3D 구조에서 실제 촬영 위치와 시점을 결정한다. 이때 특징점 정합을 위해 SIFT 특징점 추출기가 사용된다^[10].

IV. 맺음말

특징점의 검출은 영상이 갖고 있는 핵심적인 정보를 추출하고 복수의 영상 취득 센서로부터 얻은 영상들의 정렬을 위해 수행해야 할 필수적인 전처리 과정이다. 좋은 특징점 검출기는 조명과 스케일 등 영상의 취득 환경의 변화에 상관없이 핵심적인 특징점을 일관성 있게 추출해야 한다. 여러 연구 결과에 의하면 Harris 검출기가 특징점을 가장 안정적으로 검출하는 것으로 평가되고 있으며 Harris 검출기를 기본으로 스케일과 아핀 변환에 강인한 검출기로 확장이 가능하다. 스케일 변환에 강인한 Harris-Laplace 검출기는 스케일-공간 영역에서 동시에 극점값을 갖는 특징점을 찾는 것으로 이와 같은 검출기로부터 효율성을 향상시키는 방향으로 진화한 SIFT 검출기가 현재 여러 응용분야에 가장 대중적으로 활용되고 있다. 또한 검출기가 주로 실시간 처리를 요구하는 동영상 환경에서 사용되므로 처리시간을 감축하는 방향으로 진화하고 있다.

V. 감사의 글

본 논문은 국방과학연구소의 지원으로 수행되었음.

참고문헌

- [1] J. H. Fecteau and D.P. Munoz, "Saliency, relevance, and firing: a priority map for target selection," Trends in Cognitive Science, Vol.10, No.8, pp.382-390, 2006.
- [2] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," Int. J. of Computer Vision, Vol. 37, No.2, pp.151-172, 2000.
- [3] C. Harris and M. Stephens, "A combined corner and edge detector," in Alvey Vision Conference, pp.147-151, 1988.

- [4] S. M. Smith and J. M. Brady, "SUSAN . A new approach to low level image processing," International Journal of Computer Vision, Vol.23, No.34, pp.45,78, 1997.
- [5] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," Int. J. of Computer Vision, Vol.41, No.60, pp.63-86, 2004.
- [6] H. Bay, A. Ess, T. Tuytelaars, and L.V. Gool, "Speed-up robust features (SURF)," Computer Vision and Image Understanding, 110, pp.346-359, 2008.
- [7] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. Journal of Computer Vision, Vol.60. No.2, pp.91-110, 2004.
- [8] S. Branea and S. Filin, "Registration of terrestrial laser scans via image based features," ISPRS Workshop on Laser Scanning, Sep., 2007.
- [9] N. Snavely, I. Simon, M. Goesele, R. Szeliski, and M. Seitz, "Scene reconstruction and visualization from community photo collection," Proc. of IEEE, Vol.98, No.8, pp.1370-1390, 2010.
- [10] S. Bae, A. Agarwala, and F. Durand, "Computational Re-Photography," ACM Tr. on Graphics, Vol.29, No.3, 2010.



원 치 선

1978년~1982년 고려대학교 전자공학과 학사.
 1983년~University of Massachusetts, Amherst 석사 및 박사.
 1989년~1992년 금성사 선임연구원.
 1992년~현재 동국대학교 전자전기공학부 교수.
 <관심분야> 영상처리, 영상분할, 영상검색



엄 기 현

.....

1975년 서울대학교 공과대학 응용수학과 공학사.
 1977년 한국과학기술원 전산학과 이학석사.
 1994년 서울대학교 컴퓨터공학과 공학박사.
 1978년~현재 동국대학교 멀티미디어공학과 교수.
 1995년~1999년 동국대학교 정보관리처장 역임.
 2001년~2003년 동국대학교 정보산업대학 학장 역임.
 2007년 멀티미디어학회 회장
 2009년~2011년 동국대학교 영상미디어대학 학장 겸 영
 상 대학원 원장.
 <관심분야> 멀티미디어응용시스템, 멀티미디어데이터베
 이스, 게임시스템 설계



심 성 대

.....

2004년 경북대학교 전자전기공학부 공학사.
 2006년 포항공과대학교 전자전기공학과 공학석사.
 2006년~현재 국방과학연구소 연구원
 <관심분야> 국방로봇, 로봇비전, 영상처리



조 경 은

.....

1993년 동국대학교 전자계산학과 공학사.
 1995년 동국대학교 컴퓨터공학과 공학석사.
 2001년 동국대학교 컴퓨터공학과 공학박사.
 2003년~현재 동국대학교 멀티미디어공학과 부교수.
 <관심분야> 멀티미디어 정보처리, 게임 및 로봇 지능, 게
 임 알고리즘