

## 해리스 코너 포인트와 지역화를 이용한 모자이크 기반 CCTV시스템에 관한 연구

김희수<sup>†</sup>

### The Study on Mosaic-based CCTV System using Localization and Harris Corner Point

Hee-Su Kim<sup>†</sup>

#### Abstract

본 논문은 지역 부분영상을 이용한 실시간 CCTV결합 시스템에 고나한 연구로써 프레임에 관한 결합 인식방법을 제안한다. 제안된 알고리즘은 기존에 비해 결합속도를 줄이고 이에 따라 전체적인 인식률을 높이는 것을 목표로 하고 있다. 기존 방법인 SIFT알고리즘은 특허가 있고 처리속도가 늦기 때문에 본 논문에서 개선된 지역영상 재생성 방식을 이용하여 실질적으로 CCTV 처리속도에 맞추어 속도를 높였다. 논문은 지역화 이미지를 이용하여 인식률과 속도를 바탕으로 전체적인 시스템 구현과 함께 알고리즘을 소개하는 것으로 구성되어 있으며 실제 실험을 통한 성능을 비교 평가 하였다. 실시간으로 동작하는 CCTV에 적용시킴으로 각각의 화면이 아닌 Inline System으로 감시자의 신체적 피로를 줄이고 경제적인 효과를 볼 수 있고 일반 사용자들도 쉽게 사용할 수 있는 저비용 시스템을 구축하였다.

**Key words** : SIEF, FFT, Harris Corner Point

#### 1. 서 론

이미지 기반 모델링과 이미지 기반 렌더링 기술의 하나인 이미지 모자이크는 고해상도의 큰 이미지를 획득하기 위해 가장 널리 사용되는 방법이다. 간단한 이미지 모자이크 접근기법으로는 카메라 이동에 따라 달리 획득된 이미지들의 중복된 부분을 찾아 붙여서 보다 큰 이미지로 만드는 것이다. 이러한 이미지 모자이크의 방법과 응용사례는 다양하게 보고되고 있으며, 특히 최근 들어 각광을 받고있는 컴퓨터 비전과 가상 환경 분야의 중요한 기초 기술로도 활용되고 있다.<sup>[1]</sup> 특히 영상의 합성분야는 실생활에도 많이 이용되고 있다. 합성시 영상의 중첩영역의 부위를 찾는 것은 어려운 일이다. 기본적으로 특징 점을 찾는 알고리즘은 필수이고 이 찾은 영상의 분석과 인식을 위한 빠른 처리 속도 또한 고려해야할 요소이다. 기존 연구들은 크게 SIFT의

라이브어리를 이용하는 경우가 많다 이는 계산 량이 많아서 속도 면에서 느리다는 가장 큰 결점을 가지고 있다. 이를 보완하기 위한 연구들이 많이 행하여지고 있다. 그중 의 SIFT에 PCA를 이용한 방법과 Levenberg Marquardt 알고리즘을 이용한 방법 대표적인 예이다. 위 논문들은 벡터 값에 의해 인식의 결정을 짓는 것과 달리 본 논문에서는 영상의 특징을 영역에 따른 인식 방법을 사용하였다.

#### 2. 기존 방법

이미지 매칭은 다양한 분야에서 사용되어 지고 있다. 이미지 매칭을 위한 첫걸음은 특정영역을 검출해야한다는 것이다. 특정 영역을 검출하는 방법에는 점, 선, 에지, 등등이 있는데 초기에는 주로 점을 위주로 검출하였다. 점을 검출하기 위한 첫 번째 연산자는 Moravec 연산자부터 시작하게 되었는데 이는 비등방성이라는 제한이 있었기 때문에<sup>[2]</sup> Harris와 Stephens은 계산하는 비용이 많으나 검출 면에서 큰 효과를 주는 방법을 자기상관함수를 사용하여 Corner detection 을 제시하였

광주보건대학 병원전산과 (Dept. of Hospital Information Management  
Gwangju Health College)

<sup>†</sup>Corresponding author: hskim@ghc.ac.kr  
(Received : January 28, 2010, Accepted : March 20, 2010)

다. 이 Harris Corner Detection은 지금도 많이 사용하고 있으나 연산량이 많아 속도가 느리기 때문에 속도를 개선하는 방법을 고안하여<sup>[3]</sup> Susan과<sup>[4]</sup> Fast Corner Detection 등 다양한 이론들이 나오게 되었다. 이들은 noise, scale, orientation 등에 영향을 받기 때문에 특정 점 검출만 가지고 이미지 매칭을 하는데 있어 어려움을 가지고 있다. 그 중 Scale과 Orientation의 변화에도 변하지 않는 Robust 점을 검출하는<sup>[5,6]</sup> SIFT(Scale Invariant Feature Transform)가 2004년 David G. Lowe 가 최종적으로 제시한 방법으로 이미지 매칭에 있어 가장 성능이 우수하여 지금까지 로봇이나 이미지 유사판별등 많은 분야에서 사용되어지고 있다. SIFT는 그레이 스케일 이미지로 전처리 하고 각각의 가우시안 분포에 대해 이미지를 생성하고 Image Pyramid로 scale이 다른 이미지를 생성한 후, 생성된 가우시안 이미지의 차를 통하여 DOG(Difference of Gaussian)을 구한다. 이 DOG는 scale과 orientation에 강한 점들이 나오게 되는데 특징점이 많이 나오게 된다. 많은 특징 점과 고차원적인 특징 점 벡터를 사용하게 되어 매칭에서는 좋은 성과를 보여 주나 연산량이 많아 실시간 이미지 처리에는 부적절하다고 특허까지 걸려 있어 실용성에서 제약이 많다. 그래서 고안되는 방법에는 고차원 벡터를 다 낮은 차원으로 줄이는 주성분분석(PCA)을 이용한<sup>[7]</sup> PCA-SIFT나<sup>[9]</sup> Hessian 방법을 기반으로 매우 안정적인 특징점을 빠르게 찾는<sup>[8]</sup> SURF(Speeded Up Robust Features) 등이 제안되어 지금까지 여러 분야에서 사용되어지고 있다. 하지만 이것 역시 실시간 이미지 매칭에서는 약간 부족함이 있어 많은 연구가 진행되는 분야중 하나이다.

### 3. 알고리즘 순서도

본 논문에서 제안하는 프레임간의 인식시스템의 알고리즘은 CCTV 카메라에서 프레임만 추출하여 실시간으로 인식하여 재동영상화 시킨다. 그림 1은 본 논문에서 제안하는 알고리즘의 순서도를 나타낸다. 프레임의 특징부위를 추출하기위해 영상의 전처리과정이 필요하다 전처리과정은 노이즈 제거와 최적의 특징점을 찾기 위한 조건으로 사용되고 있다.

인식 부분에서는 추출된 점의 41\*41 부분적인 프레임 만들어 차연산을 통한 적절한 임계치로 최적의 점을 찾게 된다. 입력 영상이 인식을 하게 되면 인식한 부분을 최적에 맞게 한 화면으로 결합하게 된다.

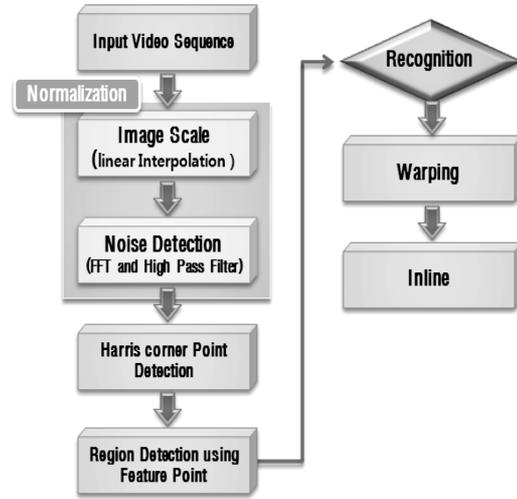


그림 1. 알고리즘 순서도.

## 4. 영상 정규화

### 4.1. 영상의 크기 변환

주파수 영역을 계산하기 쉽게 하기 위해 영상의 크기를 256\*256 크기로 설정하였다. 이는 2<sup>n</sup>을 하기위해 사용자 정의인 식으로 계단현상을 줄이기 위해 3차 회선보간법을 사용하였다. 3차 회선보간법이란 고차다항식을 이용한 보간법으로 가중치 함수를 정의 한 후 원본영상 주변 픽셀 값에 가중치를 곱한 값을 모두 합하여 픽셀 값을 계산 하는 방식을 말한다. 그림 2는 3차 회선보간법에서 사용되는 원본 픽셀들을 보여준다.

모두 16개의 픽셀을 참조하여 결과 영상픽셀의 값을 계산한다. 중앙의 점이 결과영상에서 참조해야 할 원본영상의 실수좌표라면 그 주변의 16개의 원형 점들이 3

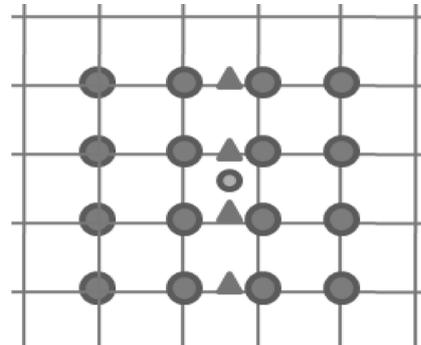


그림 2. 16포인트 원본 픽셀.

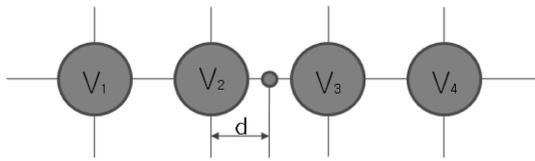


그림 3. 4포인트 원본 픽셀.

차회선 보간법에 사용되는 픽셀들을 나타낸다. 총 5번의 보간이 이용되는 16개중 4개의 삼각형위치 값을 이용하여 3차회선 보간법 수행 후 4삼각형간의 보간법 수행으로 가운데 점의 위치 값을 추정하는 것이다. 그리하여 4+1 = 5 총 5번의 보간이 이루어진다. 그림 3은 4개점으로부터 수식 1을 유도하는 그림이다.

네 좌표는  $v_1 v_2 v_3 v_4$   $v_1$ 로 표현하였고 두 번째 픽셀과 실수 좌표 사이의 거리를  $d$ 로 나타내었다  $d$ 는 0부터 1 사이의 실수 값을 가질 것이다.

이 경우  $v_1$ 에 해당하는 픽셀에서 가중치 값은  $f(1+d)$ 로 결정되며  $v_2$ 에 해당하는 픽셀의 가중치 값은  $f(d)$ 로 이와 같이 결정되는 원리를 이용하면 다음과 같은 수식 1을 구할 수 있다.

$$f(x) = \begin{cases} (a+2)|x|^3 - (a+3)|x|^2 + 1 & 0 \leq |x| \leq 1 \\ a|x|^3 - 5a|x|^2 + 8a|x| - 4a & 1 \leq |x| \leq 2 \\ 0 & 2 \leq |x| \end{cases} \quad (식1)$$

$a$ 를 여기서 -1로 사용하고 다시 쓰면 식 2와 같이 구할 수 있다.

$$f(x) = \begin{cases} |x|^3 - 2|x|^2 + 1 & 0 \leq |x| < 1 \\ -|x|^3 + 5|x|^2 - 8|x| + 4 & 1 \leq |x| < 2 \\ 0 & 2 \leq |x| \end{cases} \quad (식2)$$

이 함수를 이용하여 보간된 값  $v$ 를 계산하여 전개하면 수식 3과 같이 나타낸다. 이 식을 이용하여 크기를 변환 한다.

$$v = (-v_1 + v_2 - v_3 + v_4)d^3 + (2v_1 - 2v_2 + v_3 + v_4)d^2 + (v_1 - v_3)d + v_2 \quad (식3)$$

#### 4.2. 가우시안 고역필터

자연계에서 발생하는 모든 신호는 특정 함수들의 조합으로 표현할 수 있다. 푸리에변환이란 sin 함수와 cos 함수를 기저함수로 나타낼 수 있는 방법을 말한다. 본 논문에서는 속도를 향상시키기 위해서 일반적으로 많이 사용되는 FFT 을 사용하였다. 그림 4는 푸리에



그림 4. 푸리에 변환 영역.

변환에서 영역을 나타낸 것이다.

푸리에 변환에서는 일반적으로 픽셀의 값이 급격하게 변화하는 부분을 고주파성분이라 하는데 이는 가운데 분포하고 저저주파성분은 양 사이트에 분포하게 된다. 수식 4는 고역필터를 통과시키기 위한 수식이다.

$$H(u,v) = 1 - e^{-D^2(u,v)2D_0^2} \quad (식4)$$

수식 4에서  $D(u, v)$ 는 필터중앙 좌표에서  $(u, v)$ 좌표를 나타내고  $D_0$ 는 차단주파수이다. 차단주파수는 고역 통과 필터에서 걸러낼 주파수 성분의 양을 결정하는 값이다.

#### 4.3. 헤리스 코너 포인트 검출

영상에서 코너 포인트는 영상을 분석하는데 많이 이용되고 있다. 헤리스 코너 포인트는 기본적으로 윈도우 움직임에 따른 픽셀 값의 변화를 분석하여 결정하게 된다. 그림 5에서 첫 번째 그림은 평탄한 영역에서 윈도우의 움직임인데 윈도우가 움직여도 픽셀 값은 항상 일정하게 된다. 두 번째 그림은 경계선에 위치하는 경우이다 이는 좌우로 움직일 시에는 값의 변화가 있지만 상하로 움직이는 경우에는 값의 변화가 없다. 마지막 그림은 헤리스 코너 포인트이다 이는 좌우상하로

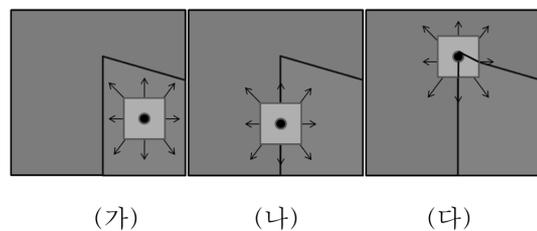


그림 5. 영역별 포인트 위치.

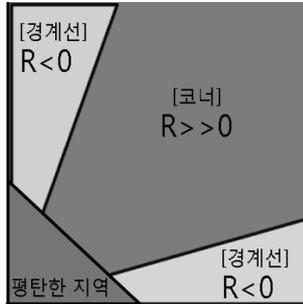


그림 6. R 값에 따른 분포.

움직여도 값의 변화가 있다 이러한 방법으로 해리스 코너 포인트를 찾는다.

이는 수식 5로 나타내보면  $\Delta x$ 와  $\Delta y$ 는 x와 y방향으로 작은 변화량을 의미하고 W는 윈도우를 의미한다.

$$c(x,y) = \sum_w [I(x,y) - I(x+\Delta x, y+\Delta y)]^2$$

$$M = \sum_{\Delta x \Delta y} w(\Delta x, \Delta y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (식5)$$

변화량을 나타내는 함수  $c(x, y)$ 는 그림 6과 같이 평탄한 영역에서는 그 값이 작게 나타나고 코너 위치는 그 값이 크게 나타나게 된다. 이는 테일러 급수에 의하여 수식 6과같이 나타낼 수 있다. 여기서 행렬 M의 고유값을  $\lambda_1, \lambda_2$ 라고 표현한다면 경계면인지 코너 포인트 인지를 그림 6과 같이 알 수 있다.

$$R = Det(M) - K \cdot Tr(M)^2$$

$$\begin{cases} Det(M) = \lambda_1 \cdot \lambda_2 \\ Tr(M) = \lambda_1 + \lambda_2 \end{cases} \quad (식6)$$

만약  $\lambda_1, \lambda_2$ 의 값이 모두 작으면 평탄하고  $\lambda_1, \lambda_2$ 중 하나는 크고 하나는 작으면 경계선이며 둘 다 큰값을 가지면 이는 코너 포인트가 되는 것이다.

4.4. 지역화(41\*41 영상) 재생성 후 인식

영상에서 나온 코너 포인트를 기준으로 41\*41 크기의 영상을 재생성하여 두 영상의 차이연산을 통하여 일정비율을 판별하는 방식으로 인식을 하였다. 그림 7은 재생성된 이미지이다.

침입자 검출에도 많이 사용되는 차이연산 식 7을 사용하였다. 여기서 92.16%의 이상의 차이정도를 나타내는 점들만을 뽑아낸다. 그 점들 중 Max&Min알고리즘

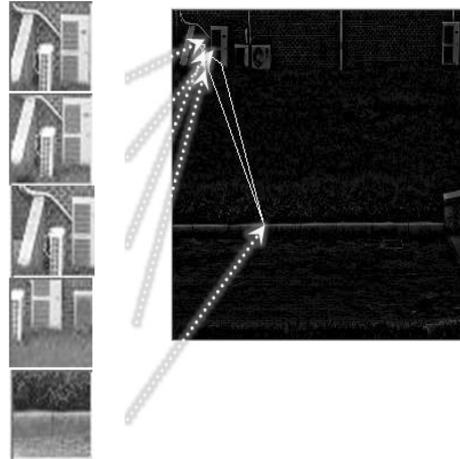


그림 7. 41\*41 이미지 생성.

을 적용하여 Max점에 변환 주기가 있는 점을 다시 뽑은후 그 전체 점의 70%이상의 점중 2점 이상을 지니고 있는 점들을 가지고 영역을 만들고 그 영역에 따른 결합을 하였다.

$$h(x, y) = |f(x, y) - g(x, y)| \quad (식7)$$

5. 실험 결과

실험은 정지영상에서 결합 정확성 테스트와 합성 속도테스트에 대해 2가지 경우로 나누어 실험 하였다. 실험환경으로는 그림 8과 같다

2가지의 영상으로 실험을 하였는데 먼저 인식을 부분이다. 인식을 보면 640\*480샘플 경우 SIFT나 Levenberg 보다 약간 떨어지는 면이 있다. 그러나 768\*1024 샘플 같은 경우에는 거의 비슷함을 그림 9를 보고 알 수 있다. 영상의 크기 조절에서 영향을 끼쳤기 때문에 나타난 현상이다.

두 번째로는 속도 면이다. SIFT는 속도가 현저히 차이가 나기에 Levenberg와 두 개만을 비교 하였다. 표 1

<p>실험 환경</p> <p>장비 : Pentium 4 3.0GHz 1G Ram</p> <p>소프트웨어 환경 : Windows XP, Visual C++6.0</p> <p>영상의 포맷 : 640 * 840 24Bit RGB 컬러영상</p> <p>768 * 1024 24Bit RGB 컬러영상</p>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------

그림 8. 실험 환경.

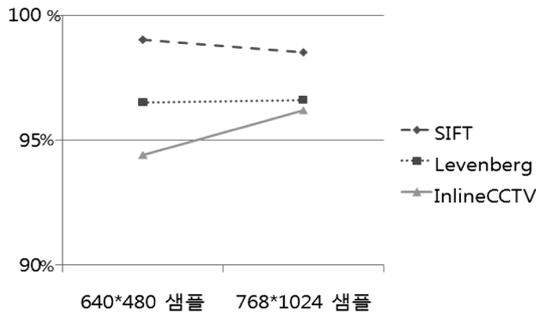


그림 9. 두 샘플 실험에 따른 그래프.

표 1. 실행 속도 비교 (단위: 초)

	Levenberg	InlineCCTV
640*480 샘플	2.395820	1.56749
768*1024 샘플	3.681138	2.12447

을 보면 Levenberg보다 약1초 가량이 빠르다는 결과를 나타냈다.

벡터를 사용하지 않고 영역인식이 인식률이 약간 떨어지지만 속도는 다른 방법보다 빠르다는 결론을 지을 수가 있다. 인식률을 높이는 게 앞으로 향후 연구계획이다.

### 6. 구성 및 응용

정보화 사회가 오면서 인간은 다양한 수단을 사용하여 위협이나 손실로부터 보호받는 상태를 만들고자 하는데 이를 보안이라 한다. 이미 인간의 보안중 일부분인 CCTV는 현재 많은 범죄를 방지 하고 유지하기에 좋은 시스템중 하나이다. CCTV는 카메라의 순간 캡처 화면을 초당 일정프레임 양으로 보여줌으로써 움직임 영상이 만들어진다. 이는 만화의 원리와 비슷하다. SIFT는 특허로 되어있는 단점과 속도가 느리다는 단점을 지니고 있다. 본 논문에서는 속도를 개선하여 실시간 CCTV에 적용하였다 일반사용자도 쉽게 사용할 수 있을 뿐만 아니라 감시자의 눈의 피로를 줄일 수 있다. 중첩 영역을 통하여 경제적인 면에서 비용을 줄일 수 있는 저비용 시스템으로 구축하였다. 그림 10은 Inline CCTV가 작동하는 그림이다.

### 7. 결 론

오늘날 CCTV의 활용을 더 효율적으로 활용하기위

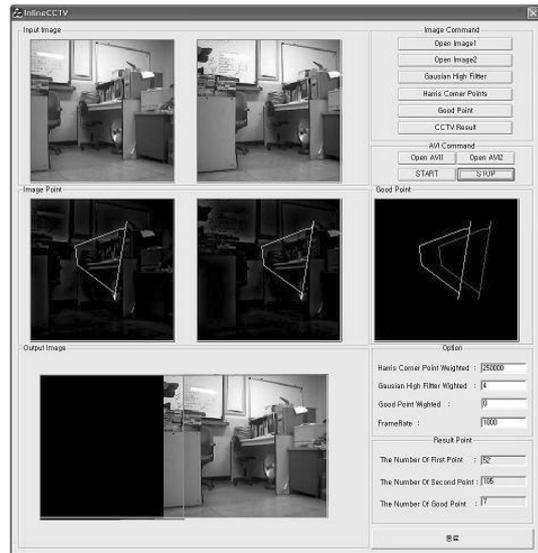


그림 10. CCTV 동작 영상.

한 시스템으로 다량의 CCTV에 InLineCCTV 기술을 이용하여 직접 CCTV를 제어하고 CCTV의 중첩 영역까지 찾아 한 라인으로 효율적으로 표현할 수 있게 만들려고 한다. 신체적으로는 감시자의 눈의 피로를 줄이고 경제적으로 비용 절감과 실용적으로 CCTV 위치 선정에도 좋은 이점이 있다. 그러나 아직까지는 인식시 필요한 주인자 선택의 개선과 영상의 회전시의 문제점에 아직 미흡하기에 이를 개선을 위해 더 많은 연구가 계속 수행해져야 한다.

### 참고문헌

- [1] Darius Burschka, Dana Cobzas, Zach Dodds, Greg Hager, Martain Jagersand, and Keith Yerec, "Recent Methods for Image-based Modeling and Rendering", IEEE Virtual Reality tutorial 1, March 2003.
- [2] C. Harris and M. j. Stephens, "A combined corner and edge detector", In Alvey Vision Conference, 1988. pp. 147-151.
- [3] S. M. Smith and M. Brady, SUSAN - A New Approach to Low Level Image Processing. *International Journal of Computer Vision*, Vol. 23(1), pp. 45-78, 1997.
- [4] M. Trajkovic and M. Hedley, Fast Corner Detection. *Image and Vision Computing*, Vol. 16(2), pp. 75-87, 1998.
- [5] David G. Lowe, "Object Recognition from Local Scale - Invariant Features". Proc of the 7th Int.

- Conf. on Computer Vision, pp. 1150-1157, Corfu, Greece, 1999.
- [6] David G. Lowe, "Distinctive image features from scale-invariant keypoints", Accepted for publication Int. Journal of Computer Vision, 2004.
- [7] Y. Ke and R. Sukthankar, PCA-SIFT: A more distinctive representation for local image descriptors. In: CVPR. 506-513. 2004.
- [8] Herbert Bay Tinne Tuytelaars, and Luc Van Gool. "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding, Volume 110 Issue 3, June 2008.
- [9] K. Mikolajczyk and C. Schmid, Indexing based on scale invariant interest points. In:ICCV. Volume 1. 525-531. 2001.
- [10] A. J. Lacey, N. Pinitkarn and N. A. Thacker, "An evaluation of the performance of RANSAC algorithms for stereo camera calibration", The 11th British Machine Vision Conference, Bristol, UK, 2000.