듀얼 모드(고정형+PTZ 카메라) 감시 카메라를 이용한 효과적인 화상 감시 시스템에 관한 연구

Study on Effective Visual Surveillance System using Dual-mode(Fixed+Pan/Tilt/Zoom) Camera

김 기 석*, 이 삭, 박 종 섭, 조 재 수
 (Giseok Kim¹, Saac Lee¹, Jong-Seop Park¹, and Jae-Soo Cho¹)

¹Korea University of Technology and Education

Abstract: An effective dual-mode camera system(a passive wide-angle camera and a pan-tilt-zoom camera) is proposed in order to improve the performance of visual surveillance. The fixed wide-angle camera is used to monitor large open areas, but the moving objects on the images are too small to view in detail. And, the PTZ camera is capable of increasing the monitoring area and enhancing the image quality by tracking and zooming in on a specific moving target. However, its FOV (Field of View) is limited when zooming in on a specific target. Therefore, the cooperation of wide-angle and PTZ cameras is complementary. In this paper, we propose an automatic initial set-up algorithm and coordinate transform method from the wide-angle camera coordinate to the PTZ one, which are necessary to achieve the cooperation. The automatic initial set-up algorithm is able to synchronize the views of two cameras. When a moving object appears on the image plane of a wide-angle camera after the initial set-up positioning, the obtained values of the wide-angle camera should be transformed to the PTZ values based on the coordinate transform method. We also develop the PTZ control method. Various in-door and out-door experiments show that the proposed dual-camera system is feasible for the effective visual surveillance.

Keywords: dual-camera system, visual surveillance system, moving object tracking

1. 서론

지능형 영상감시 분야에서 한 대의 고정형 카메라만을 이용해서 넓은 영역을 감시하는 경우, 매우 먼 거리에서 감시하게 되어 실제 감시 대상 물체에 대한 변별력이 떨어지는 문제점이 있다. 또한 PTZ 카메라 한 대만을 이용하는 경우, 하나의 물체를 추적하는 동안 다른 영역을 감시할 수 없고, PTZ 카메라를 움직이면서 계속해서 물체를 추적하는 추적 알고리즘의 성능도 실용화하기에는 정확도가 부족한 상황이다. 반면에 고정 카메라와 PTZ 카메라를 동시에 활용할 경우 안전 영역 전체에 대한 감시와 특정 위험 요소에 대한 정밀 감시가 동시에 이뤄질 수 있기 때문에 감시 목적에 있어서 매우 효율적이다. 고정된 카메라에서의 움직이는 물체 추적 알고리즘의 경우 비교적 실용적으로 사용할 수 있는 수준이다.

그림 1은 본 논문에서 제안하는 듀얼 모드 카메라를 활용한 화상 감시 시스템의 구성도를 개념적으로 잘 보여주고 있다. 고정 카메라는 감시 영역의 전체를 관찰 한다는 의미로 Global Camera라고 정의할 수 있고, PTZ 카메라의 경우 움직이는 물체의 세세한 부분을 감시한다는 의미로

Local Camera라고 정의할 수 있다. 듀얼 카메라 감시 시스템에서 핵심이 되는 부분은 고정 카메라로 입력된 영상에서 움직이는 물체를 추적하여 추적된 목표 물체에 대해 PTZ 카메라로 정확히 영상 중심으로 이동시키고(Pan/Tilt 제어) 영상의 변별력을 위해 그 크기를 확대(zooming)하는 것이다. 이를 위해 이전 연구[1]에서는 두 카메라간의 특성을 기하학적으로 분석하여 정확한 확대가 이루어지도록 하고 있다. 그러나 이렇게 하기 위해서 부가적으로 추적하고자 하는 물체에 대한 절대적인 거리를 먼저 측정해야 하는 문제점이 있다. 물론 두 대의 카메라를 이용하면 물체의 거

그림 1. 듀얼 모드 카메라(고정형 카메라와 PTZ 카메라)를 활용한 화상 감시 시스템의 구성도.

Fig. 1. Visual Surveillance System with Dual-mode cameras(static and PTZ camera).

* 책임저자(Corresponding Author)

논문접수: 2012. 2. 13., 수정: 2012. 5. 7., 채택확정: 2012. 5. 23.

김기석, 이삭: 한국기술교육대학교 대학원 컴퓨터공학과

ladofa@koreatech.ac.kr/isaac@koreatech.ac.kr)

박종섭, 조재수: 한국기술교육대학교 컴퓨터공학부

(pjs0329s@koreatech.ac.kr/jaesoo27@koreatech.ac.kr)

리를 측정할 수는 있지만, 두 카메라의 위치에 따른 거리 추정의 한계가 있고, 처리 시간 및 정확도 측면에서 구체적으로 구현하기에는 현실적으로 많은 어려움이 있다. 또 다른 이전연구[2]에서는 고정 카메라와 PTZ 카메라의 조합이 아닌 두 대의 PTZ 카메라를 활용하여 감시 성능을 높이는 연구가 발표되기도 하였다. 일반적으로 PTZ 카메라는 고정 카메라에 비하여 비용이 작게는 5배에서 15배의 비용 차이가 있고, 하나의 감시 영역에 두 대의 고가의 PTZ 카메라를 설치하는 것은 비효율적이고, 현실적이지 않은 문제점이 있다.

본 연구에서는 고정 카메라와 PTZ 카메라를 한 셋트로 하는 듀얼 카메라 시스템에 관한 연구로서 PTZ 카메라와 고정 카메라가 같은 위치에 있다고 가정한다. 이를 위해서는 PTZ 카메라와 고정 카메라의 실제 위치를 최대한 가깝게 설치할 필요가 있고, 두 카메라의 감시 뷰가 최대한 일치되도록 설정할 필요가 있다. 하지만, 실제로 PTZ 카메라와 고정 카메라가 똑같은 위치에 있지 않고, PTZ 카메라 제어로는 보정할 수 없는 회전까지 고려해야 하기 때문에 아무리 수동적으로 두 카메라의 뷰를 일치시키는 것은 한계가 있고, 설치하기도 어려운 문제가 있기 때문에 두 카메라의 뷰를 일치시키는 자동 설치 알고리즘을 포함하고 있다.

II 장에서는 기존의 움직이는 물체 추적 알고리즘 적용에 대한 문제점을 간단하게 분석하고, III 장에서는 제안하는 듀얼 모드 화상 감시 시스템을 설명한다. IV 장에서는 제안된 듀얼 모드 화상 감시 시스템에 대한 실내외 실험 결과를 보이고, 그 효용성을 입증한다. 마지막으로 V 장에서는 결론 및 추후 연구 과제에 대하여 기술한다.

II. 움직이는 물체 추적 알고리즘 적용에 대한 문제점 분석

1. 기존 화상 감시 시스템의 문제점 분석

아날로그 CCTV와 디지털 기술을 이용한 최근의 디지털 비디오 레코더 모두에 있어서 가장 큰 문제점이라고 하면 실제 범죄가 발생하였을 때 감시된 화면(저장되어 있는 화면)이 그 범죄자를 잘 식별할 수 없다는 것이다. 그림 2(a)는 2009년도 인천에서 발생한 어린이 유괴사건 용의자 화면이고, (b)는 2007년 발생한 ‘안산역 토막 살인 사건’ 용의자의 모습이 담긴 감시카메라 동영상 화면 중 범인이 포착된 가장 변별력이 있다고 하는 화면이다. 그림에서 보는 바와 같이 비록 감시카메라에 촬영된 범인의 모습임에도 불구하고, 범인의 얼굴이 변별력이 많이 떨어지는 문제점이 있다. 그 이유는 고정된 카메라에서 감시 영역을 넓게 커버하기 위해서는 화각(field-of view)을 매우 넓게 유지할 수밖에 없기 때문이다. 현재 전 세계적으로 대부분의 화상 감시 시스템이 이러한 방식으로 설치 및 운영되고 있다.

2. 기존 화상 감시 시스템에 움직이는 물체 추적 알고리즘 적용에 대한 문제점 분석

영상을 이용한 움직이는 물체 자동추적 기술[4]은 가장 먼저 군사용 무기에 응용되면서 그 기술이 발전하여 왔지만, 최근에는 민수용 무인 화상 감시 시스템(unmanned visual surveillance system), 지능형 교통정보 시스템, 지능형

(a) A kidnapper suspect

(b) A murderer suspect on evidence images.

그림 2. 사건발생 시 촬영된 용의자 화면 예들.

Fig. 2. Examples of scene of incident.

로봇 등에 부분적으로 활용되고 있는 첨단 영상신호처리 기술이다. 하지만, 다양한 물체 추적 알고리즘이 연구되고 발표되고 있지만, 아직까지 이러한 물체 추적 알고리즘 또는 추적 카메라가 실제 화상 감시 시스템에서의 활용은 매우 미미한 실정이다.

카메라가 고정되어 있는 경우에 대한 기존의 국내외 많은 물체 추적 관련 연구들(차영상법, 디스터번스맵, optical flow, background modeling 및 subtraction)은 카메라의 움직임이 없는 상태(고정 카메라)에서의 영상을 이용하여 움직임 정보를 추출하는 방법으로 연구되어왔다. 즉, 카메라의 움직임은 카메라 움직임을 전역보상하면 된다고 가정하고 있지만, 실제 화상 감시 시스템에서 카메라의 전체 움직임을 보상하면서, 움직이는 물체를 추적함과 동시에 입력 영상을 압축해서 저장장치에 기록하기에는 너무나 많은 계산량을 필요로 하기 때문에 추적성능을 떠나서 실제 구현성이 매우 떨어지는 문제점이 있다. 즉, 카메라의 전역 움직임을 고려하지 않은 많은 추적기법들은 실제 카메라가 능동적으로 움직이는 환경에서는 적용하기 힘든 문제점이 있다. 그리고 추적성능이 비교적 좋은 다양한 특징점 또는 특징모델을 이용한 추적기법들의 경우 계산량이 너무 많아 실시간 구현이 어렵고, 움직이는 물체의 변별력을 높이기 위해서는 반드시 줌인(zoom in) 제어가 필요하다. 하지만 현재까지 알려진 추적 알고리즘 중에 줌인 제어까지 할 수 있는 추적 알고리즘은 거의 없다고 할 수 있다.

추적 알고리즘의 성능에 대한 문제만 아니라 근본적으로 움직이는 물체 추적 감시 카메라가 현재의 화상 감시 시스템(CCTV)에 적용될 수 없는 또 다른 큰 이유는 감시영상을 저장하는 디지털 비디오 레코더(DVR)의 동영상 압축방법에 있다. 카메라가 자동으로 움직일 경우, 입력 영상에서

움직이는 물체에 대한 변별력은 크게 증가하겠지만, 카메라의 움직임으로 인한 추가적인 압축 데이터량이 크게 증가하게 되어(MPEG-2, 4 및 H.264 등 압축알고리즘의 원리상), 카메라를 자동으로 움직이게 할 경우 저장되는 데이터량이 매우 많아지는 문제점이 있다.

이러한 내용을 기초로 본 논문에서는 기존의 화상 감시 시스템의 성능을 높이고, 움직이는 물체 추적 기술을 더욱 효과적으로 응용할 수 있도록 하는 듀얼 모드 감시 카메라 시스템을 제안한다.

III. 듀얼 모드 감시 카메라를 이용한 효과적인 화상 감시 시스템

제안하는 듀얼 모드 감시 카메라를 이용한 효과적인 화상 감시 시스템의 전체 순서도는 그림 3(a)에 나타내었다. 듀얼 모드 카메라를 이용한 화상 감시 시스템에서 감시기능을 높이기 위해서는 먼저 각각의 카메라의 역할 정의가 필요하다. 고정된 카메라에서는 일반적인 화상 감시 시스템의 카메라 역할을 수행하면서 입력 영상을 통하여 움직이는 물체 정보(중심좌표(x, y))와 움직임 물체 크기정보(s))를 획득하는 역할을 한다. 고정된 카메라에서 획득된 움직임은 물체 정보($G(x, y, s)$)는 PTZ 카메라 영상 좌표로($P(x, y, s)$) 변환된 후, 움직이는 물체를 계속해서 추적하기 위해 움직이는 물체를 화면의 중심으로 팬 및 틸트 제어를 수행하고, 물체의 변별력을 높이기 위해 배율을 높이는 줌 제어를 수행하게 된다. PTZ 카메라 제어의 목적은 최대한 움직이는 물체를 오래도록 추적하면서, 추적물체의 변별력을 높이기 위해 최대한 감시화면에 적당한 배율로 모니터링하면서, 변별력있는 감시물체를 HDD에 저장하는 것이다. 고정 카메라는 기존의 화상 감시 시스템처럼 동영상으로 저장되고, PTZ 카메라는 변별력있는 장면만(사람의 얼굴 및 자동차의

번호판 장면) 저장한다.

이러한 감시 시스템을 위해서는 먼저 고정 카메라와 PTZ 카메라의 기본위치 자동 설정 알고리즘이 필요하고, 두 카메라간 좌표 변환 알고리즘 및 PTZ 카메라를 Pan/Tilt/Zoom 제어하는 방법이 필요하다. 카메라는 기본적으로 [9]에 의해 보정되어 있는 상태이다.

1. 고정 카메라와 PTZ 카메라의 기본위치 자동 설정 방법

고정 카메라와 PTZ 카메라의 기본위치 자동 설정 알고리즘의 순서도는 그림 3(b)와 같이 정리할 수 있다[5]. 기본적으로 고정 카메라를 먼저 설치한 후, 최대한 고정 카메라와 가까운 위치에 PTZ 카메라를 수동 설치하고, PTZ 카메라 뷰도 고정 카메라와 어느 정도 일치할 수 있도록 수동으로 팬/틸트/줌의 위치를 설정한다. 그 다음은 SURF 알고리즘[6]을 이용하여 고정 카메라 영상의 특징점과 PTZ 카메라 영상의 특징점을 매칭시켜서 호모그래피 변환(homography transform)을 통해 카메라간 기하학적 차이를 인식한 후, PTZ 카메라를 제어(Pan, Tilt, Zoom)한다. PTZ 카메라의 위치를 이동해 가면서 PTZ 카메라의 위치가 수렴할 때 까지 PTZ 카메라의 초기 위치를 자동으로 제어하여 두 카메라를 자동 위치시킨다. 즉 고정형 카메라의 영상을 레퍼런스로 해서 PTZ 카메라에서의 호모그래피 행렬을 구하고, 결과를 토대로 PTZ 카메라를 직접 움직이고 고정형 카메라와 중심이 맞도록 이동하는 과정을 반복 수행하면, 두 특징점 쌍의 좌표가 거의 일치하는 방향으로 수렴하게 되어, 고정형 카메라와 PTZ 카메라의 중심이 서로 일치하는 결과를 얻을 수 있다. 다른 매칭 방법으로서 [10,11]의 알고리즘을 응용할 수도 있다.

그림 4는 두 영상간 성공적인 정합이 이루어진 경우와 정합이 실패한 경우의 예시를 보여주고 있다. 화면 상에 전혀 겹치는 영역이 없으면 정합이 원론적으로 불가능하므로 어느 정도 겹치는 영역이 필요하다(실험적으로 약 1/4 이상). 이를 위해서는 초기 수동설치 시 최소 두 카메라 영상이 1/4 이상 겹치게 설치하는 것이 필요하거나 자동으로 감시 영역 전체를 미리 탐색하는 방법을 통하여 겹치는 부분을 찾는 과정이 부가적으로 필요할 수도 있다.

(a) Flow chart of the purposed system. (b) Auto position set-up algorithm for the PTZ camera.

그림 3. 제안하는 듀얼 모드 화상 감시 시스템.

Fig. 3. Proposed dual-mode visual surveillance system.

(a) Succeed case.

(b) Failure case.

그림 4. 특징점 Descriptor 정합 결과.

Fig. 4. The descriptor matching results.

그림 5. 고정 카메라와 PTZ 카메라의 좌표 변환.

Fig. 5. Coordinate transformation from the static camera image to the PTZ camera image.

2. 고정 카메라와 PTZ 카메라간의 좌표 변환 알고리즘

PTZ 카메라로 움직이는 물체를 추적하기 위해서는 고정 카메라에서 수집된 움직이는 물체 정보를 그림 5와 같이 고정 카메라에 나타난 물체의 특정 이미지 좌표($G(x, y)$)가 PTZ 카메라에서 나타나는 영상 좌표 $P(x, y)$ 로 변환해야 한다. 실제로 좌표 변환 알고리즘을 통하여 변환된 좌표 $\tilde{P}(x, y)$ 는 실제 PTZ 카메라상의 좌표 $P(x, y)$ 와 차이가 날 수 있다. 좋은 성능의 좌표매핑 시스템일수록 전체 영역에서 오차가 작을 것이다. 본 논문에서는 두 카메라의 좌표 변환을 위하여 들로네 삼각화[7]를 이용한 좌표 변환과 호모그래피 행렬을 이용한 좌표 변환 두 방법을 제안한다.

2.1 호모그래피 행렬을 이용한 좌표매핑

다수의 특징점 정합을 통해 생성된 특징점 정합 쌍들을 이용하여 두 카메라 시스템간의 호모그래피 변환 행렬 식 (1)과 같이 표현할 수 있다. 변환 행렬을 구하는 과정은 식 (2)에서 행렬 H의 해를 구하는 과정과 동일하다.

$$\begin{bmatrix} w\tilde{x}_i \\ w\tilde{y}_i \\ w \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (1)$$

식 (1)의 호모그래피 행렬 파라미터를 계산하기 위해 N 개의 매칭된 특징점들을 활용하여 식 (2)와 같이 호모그래피 행렬을 계산할 수 있다.

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -\tilde{x}_1 x_1 & -\tilde{x}_1 y_1 & -\tilde{x}_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -\tilde{y}_1 x_1 & -\tilde{y}_1 y_1 & -\tilde{y}_1 \\ : & : & : & : & : & : & : & : & : \\ x_N & y_N & 1 & 0 & 0 & 0 & -\tilde{x}_N x_N & -\tilde{x}_N y_N & -\tilde{x}_N \\ 0 & 0 & 0 & x_N & y_N & 1 & -\tilde{y}_N x_N & -\tilde{y}_N y_N & -\tilde{y}_N \end{bmatrix} \begin{bmatrix} H_{11} \\ H_{12} \\ H_{13} \\ H_{21} \\ H_{22} \\ H_{23} \\ H_{31} \\ H_{32} \\ H_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ : \\ 0 \\ 0 \end{bmatrix} \leftrightarrow Ah = B \quad (2)$$

여기서 $((x_1, y_1), (x_2, y_2), \dots, (x_N, y_N))$ 은 고정 카메라에서 검출된 N개의 특징점 좌표이고, $((\tilde{x}_1, \tilde{y}_1), (\tilde{x}_2, \tilde{y}_2), \dots, (\tilde{x}_N, \tilde{y}_N))$ 은 PTZ 카메라에서 매칭된 특징점 좌표이다. 특징점은 SURF [6] 알고리즘을 이용하였다. 최소 10개 이상의 특징점 쌍을 구하면 호모그래피 행렬의 모든 미지수를 구할 수 있다.

여기에서 나타난 이동, 스케일, 회전을 상쇄시킬 수 있도록 PTZ 카메라를 움직이고, 다시 이 과정을 반복한다. 호모그래피 변환의 평행 이동과 스케일은 PTZ 카메라의 물리적 이동(팬, 틸트, 줌)으로 상쇄할 수 있다. 회전의 경우에는

입력 영상에 대한 회전 변환으로 보정이 가능하다. PTZ 카메라의 뷰가 고정형 카메라와 거의 일치하게 되면 그 때부터는 일정한 평균을 중심으로 수렴하게 된다. 시간 t에 대한 제어 입력값 x_t 가 주어졌을 때, x_t 가 수렴할 때의 시간 t^c 는 이전 N개의 입력 동안의 표준 편차가 가장 적을 때로 한다. 이는 식 (3)으로 표현될 수 있고, 식 (3)의 수렴된 값 x^c 는 수렴 시점으로부터 N 기간 동안의 평균으로 식 (4)와 같이 표현할 수 있다.

$$t^c = \arg\text{Min}_t(\text{dev}(x_{t-N} \dots x_t)) \quad (3)$$

$$x^c = E(x_{t-N} \dots x_t) \quad (4)$$

여기서 $\text{dev}(\dots)$ 는 N기간 동안의 표준편차를 의미하고, $E(\dots)$ 는 평균값을 의미한다.

2.2 들로네 삼각화를 이용한 좌표 변환

들로네 삼각화(delaunay triangulation) [7]란 공간상의 점들을 연결하여 삼각형들의 집합으로 구성하는 삼각화 기법 중 하나으로써, 삼각형들의 최소 내각의 크기가 최대가 되도록 삼각형을 구성한다. 하나의 삼각형에 외접하는 외접원 내부에는 다른 점들이 포함되지 않는다. 각각의 고정 카메라의 특징점에 대해 맵핑되는 PTZ 카메라의 특징점들을 들로네 삼각화로 구한 영역별로 평면 방정식을 구함으로써, 맵핑이 이루어지지 않은 영역에 대해서 어느 정도 체계화된 좌표 변환 값을 얻을 수 있게 된다. 그림 6은 PTZ 카메라 영상내의 특징점들을 들로네 삼각화 한 결과를 보여주고 있다.

3. PTZ 카메라 제어방법

PTZ 카메라 제어의 목적은 추적하고자 하는 물체의 변별력을 높이기 위해 추적하고 있는 물체가 PTZ 카메라 영상의 중심에 오도록 하고, 추적중인 물체의 크기도 줌(zoom) 신호를 통하여 확대할 수 있도록 하는 것이다.

3.1 수학모델에 의한 팬/틸트 제어

팬과 틸트 제어신호의 값은 추적중인 물체의 중심이 영상의 중심에서 얼마나 멀어졌는가를 각도로 계산하여 제어 신호를 발생한다. 여기서 물체의 크기 혹은 물체와 카메라

그림 6. PTZ 카메라 영상내의 특징점들을 들로네 삼각화 한 결과.

Fig. 6. A result of Delaunay Triangulation of feature points on the PTZ camera.

그림 7. 카메라 화각과 영상의 관계도.

Fig. 7. The relation of angle of view and image.

간의 거리는 아무런 상관이 없고 오직 물체 중심의 이동 거리만으로 각도를 구할 수 있다. 그림 7(a)에서 음영으로 표시된 것이 추적 대상 물체이다. 영상중심으로부터 물체의 중심이 떨어진 거리 x 가 주어질 때 팬 각도 θ 는 이미지상의 소실점(카메라의 위치)을 만들어 삼각함수를 통해 구할 수 있다. 그림 7(b)에서 보는 바와 같이 카메라가 가장 끝을 가리킬 때의 팬 각도 θ_0 , 영상의 크기 w_0 , 물체의 중심이 영상의 중심으로부터 떨어진 거리 x 가 주어졌을 때는 식 (5)와 같다.

$$\theta = \tan^{-1}\left(\frac{x}{d}\right), \quad d = \frac{w_0}{2\tan(\theta_0)} \quad (5)$$

여기서 d 는 카메라의 초점 거리로서 렌즈와 CCD 센서간의 물리적인 거리가 아닌 화소 단위의 이미지상의 거리이다. 물체가 화면 끝으로 이동했을 때의 각도 θ_0 는 PTZ 카메라의 화각에 의해 미리 주어진다. 틸트 제어에 대해서도 팬 제어와 같은 방법을 적용할 수 있다.

3.2 줌(zoom) 제어

줌 제어의 경우, 전체 영상화면의 크기에 비해서 추적 물체의 크기가 차지하는 비율을 통해 줌에 적용할 배율을 정한다. 추적 물체의 가로, 세로의 비율과 영상의 가로 세로의 비율이 다르기 때문에 그 중에서 더 적은 배율을 적용한다. 이것을 식으로 표현하면 식 (6)과 식 (7)로 표현할 수 있다.

$$r = r_0 \cdot \min(r_w, r_h) \quad (6)$$

$$r_w = \frac{w_0}{w}, \quad r_h = \frac{h_0}{h} \quad (7)$$

여기서 w, h 는 현재 추적중인 물체의 가로, 세로의 길이를 의미하고, w_0, h_0 는 영상화면의 가로와 세로의 크기를 나타낸다. 즉, 영상의 w_0, h_0 에 대한 물체의 배율이 r_0 이고, 여기서 가로 또는 세로 확대배율 중 더 적은 배율만큼 확대 배율이 새롭게 제어하기 위한 줌 배율이 된다. 실제 배율 적용은 움직이는 물체가 PTZ 뷰를 벗어나지 않도록 약간 적은 배율을 적용하는 것이 여러 실험을 통하여 경험적으로 알 수 있었다.

4. 고정형 카메라의 물체 추적

직접 추적 대상 물체에 대한 테스트를 위해 이전물체 추적연구[8]에서 밝힌 배경 분리 알고리즘을 응용했다. 배경

분리 후, 열림과 닫힘 연산을 이용해 노이즈를 제거하고 라벨링을 통해 특정 물체 영역을 얻어 낸다. 본 논문에서는 이전연구[8]의 방법에 영상 노이즈를 제거하기 위한 개념을 추가했고, 연산 시간을 단축하기 위한 근사 방법을 새로 제안한다.

영상내 특정 픽셀 위치 x 에 대해 입력 영상 $I(x)$ 에 대한 배경 판정 함수 $F(x)$ 는 다음과 같다.

$$F(x) = \begin{cases} 1 & \text{if } I(x) < m(x) - (d(x) + b) \vee \\ & I(x) > n(x) + (d(x) + b) \\ 0 & \text{else} \end{cases} \quad (8)$$

여기서 $d(x)$ 는 특정 화소의 배경 변화에 대한 노이즈이며 일정 기간 동안의 분산으로 구한다. 상수 b 는 일종의 바이어스로서 영상 노이즈가 가지는 분산이다. $m(x)$ 는 일정 기간 동안의 $I(x)$ 에 대한 최소값, $n(x)$ 는 최대값으로서 일정 기간 동안의 각 화소값에 대한 통계로 구한다.

매 프레임마다 모든 화소에 대한 통계를 내기 위해서는 정해진 기간 전체의 영상을 모두 저장하고 있어야 하고, 통계에 대한 연산과정이 오래 걸리기 때문에 가중치 합의 방법을 통해 근사적으로 다음 식 (9)와 같이 구한다.

$$\begin{aligned} m(x) &= \begin{cases} m(x) & \text{if } m(x) < I(x) \\ \alpha \cdot m(x) + (1 - \alpha) \cdot I(x) & \text{else} \end{cases} \\ \text{average} &= \alpha \cdot \text{average} + (1 - \alpha) \cdot I(x) \\ d(x) &= \alpha \cdot d(x) + (1 - \alpha) \cdot |I(x) - \text{average}| \end{aligned} \quad (9)$$

여기서 α 는 (0, 1)사이의 상수값으로 배경을 학습하는 속도를 결정한다.

IV. 실험 결과 및 분석

그림 8은 실험에 사용된 두 카메라가 설치된 모습을 보여주고 있다. 실내 실험을 위해서는 실험실의 천장에 두 카메라를 설치하였고, 실외 실험은 이동이 가능하게 설치한 후, 건물 옥상에서 실험하였다. 제한된 감시 시스템에 대한 실험을 위해서 PTZ 카메라의 경우 AXIS 232D+를 사용했다.

1. 두 카메라간 자동 설정 알고리즘 성능 실험

PTZ 카메라는 무작위의 위치에서 시작했으며 고정형 카메라는 특징점이 많은 곳과 특징점이 거의 없는 벽면에 큰 배율(렌즈로 조정)을 적용한 두 장면에 대해서 각각 20회 실험했다. 각 카메라의 입력 영상에 대해서는 카메라 왜곡 보정(internal calibration)을 수행한 뒤 본 논문에서 제시한 방법을 적용했다. 실험에 쓰인 레퍼런스가 되는 고정형 카메라의 영상은 그림 9에 나타내었다.

그림 8. 실험에 사용된 두 카메라가 설치된 모습.

Fig. 8. The installed cameras.

(a) A view with many feature points. (b) Another with a little feature points.

그림 9. 실험에 쓰인 고정형 카메라의 레퍼런스 영상.

Fig. 9. The views of static camera for experiments

그림 10. PTZ 카메라의 자동 위치 설정에 따른 중심 이동.

Fig. 10. A control sequence of the PTZ camera auto set-up.

표 1. PTZ 카메라 자동 설정 실험 결과.

Table 1. A results of PTZ camera auto set-up.

	그림 9(a) 실험	그림 9(b) 실험
평균 반복 횟수(t^c)	9.24	13.20
평균 매칭 특징점 수/ 총 특징점 수	173.7/1732.4	39.2/373.0

실험 결과는 그림 10과 표 1에 나타내었다. 평균 반복 횟수는 특징점이 적은 레퍼런스의 경우 수렴하는데 더 오래 걸리는 것으로 나타났다. 실질적인 수렴 지점까지 카메라가 이동하는 데는 3, 4회로 모두 동일했지만 매칭되는 특징점의 수가 적은 경우 최소값으로 도달하는데 더 많은 반복이 필요했다. 식 (3)에서의 N 값을 5로 했기 때문에 최대 4~8번의 반복 사이에 수렴 지점에 도달한 것으로 분석된다. 한편 평균 매칭 특징점 수는 레퍼런스가 되는 고정형 카메라와 PTZ카메라의 SURF 특징점 매칭에 대해 매칭된 개수/고정형 카메라의 총 특징점 개수를 나타낸 것이다. 이는 N 프레임의 수렴 기간에 포함되는 영상만을 통계로 했다.

실험 결과 총 40회 모두 성공적인 보정 결과를 보여주었다. 특히 그림 9(b)와 같이 제한된 특징점을 가지는 환경에서도 PTZ 카메라가 수렴하는 좋은 결과를 나타냈다.

2. 두 카메라 좌표 변환 알고리즘 성능 실험

두 카메라 좌표 변환 알고리즘의 성능을 실험하기 위해 고정 카메라 좌표를 PTZ 좌표로 변환된 후 그 오차가 어느 정도 되는지에 대한 성능을 실험하였다. 여기서 오차(error)

(a) (b)

그림 11. 좌표 변환 알고리즘 테스트 결과 화면, (a) 고정 카메라에서 선택된 점들, (b) PTZ 카메라에서 제안된 알고리즘에 의해 좌표 변환된 점들(파란점: 사람이 눈으로 직접 선택, 빨간점: 들로네 삼각화 방법, 초록점: 호모그래피 행렬 방법).

Fig. 11. A result of coordinate transformation from (a) manual selected points on static camera to (b) blue manual truths, red points by the Delaunay Triangulation method, and green points by the homography matrix method on PTZ camera.

는 실제 변환되어야 하는 점(true pixel point)과 좌표 변환 알고리즘에 의하여 예측된 점(estimated pixel point) 간의 유클리드 거리를 뜻하며 두 점의 오차는 다음 식 (10)과 같이 정의할 수 있다.

$$\text{오차}(\text{error}) = \sqrt{(x - \hat{x})^2 + (y - \hat{y})^2} \quad (10)$$

여기서 (x, y) 는 실제 좌표이고, $(\hat{x}, \hat{y})$ 은 좌표 변환 알고리즘에 의해 예측된 점의 좌표를 의미한다. 실제 좌표는 직접 사람이 눈으로 확인해 결정하였다. 눈으로 확인된 지점과 알고리즘으로 예측된 지점간의 오차를 통해 제안된 좌표 변환 알고리즘의 정확도를 판단했다. 실험은 고정 카메라의 입력 영상 평면에 대해 임의의 위치 100곳을 선택하여 실험하였다. 실제 매핑되어야 하는 위치는 육안으로 확인하였을 때 고정 카메라 영상 좌표와 PTZ 영상 좌표가 동일한 지점을 표시하는 것으로 보이는 좌표를 직접 선택하였다.

그림 11(a)는 고정 카메라에서 성능실험을 위해 선택한 지점들을 나타내고, 그림 11(b)는 PTZ 카메라에서 제안된 두 개의 알고리즘에 의해 좌표 변환된 점들(파란점: 사람이 눈으로 선택, 빨간점: 들로네 삼각화 방법, 초록점: 호모그래피 행렬 방법).

(a) (b)

그림 12. 좌표 변환 오차, (a) 들로네 삼각화를 이용했을 경우 좌표 변환 오차, (b) 호모그래피 방법을 사용했을 때의 좌표 변환 오차.

Fig. 12. A errors of transformation by (a) the Delaunay Triangulation method and (b) the homography matrix method.

래피 행렬 방법)을 잘 보여주고 있다. 그리고 그림 12는 각각 들로네 삼각화를 통한 방법과 호모그래피 행렬을 이용한 방법에 대한 변환오차를 3차원으로 표현한 것이다. 그래프의 X축과 Y축은 각각 맵핑된 점의 X좌표와 Y좌표를 나타내며, Z축은 오차를 의미한다. 즉 바닥 평면이 영상 평면이며 높이가 오차를 나타낸다.

실험 결과를 분석하면 두 방법 모두 실제 맵핑되어야 할 좌표에 근사한 결과를 보였으나 전체적으로 들로네 삼각화를 사용했을 때의 오차가 더 작은 경향을 보인다. 좌표가 영상의 모서리 부분일 경우에는 오차가 상당히 증가하는 것을 관찰할 수 있는데, 이는 두 카메라간의 영상 차에 따른 영향이 크다. 이러한 오차는 화면 외각 부분이 아니더라도 물체와 카메라 간의 거리가 달라짐에 따라 발생할 수 있다.

표 2는 두 가지 좌표 변환 방법에 대한 영상의 영역별 성능비교 결과를 정리한 것이다. 안쪽 영역은 영상 전체 면적의 1/2를 차지하는 중심 영역이며 바깥쪽은 이를 제외한 영역이다. 이 결과에서 알 수 있듯이 들로네 삼각화를 이용한 선형 보간 방법이 호모그래피 방법보다는 안쪽 영역 및 전체적으로 더 좋은 성능을 보였으며, 실제 응용에 있어서는 두 알고리즘 모두 실용적인 수준의 오차를 나타냈다. 적용시키려는 환경에 따라 적절한 알고리즘을 선택할 수 있을

표 2. 두 좌표 변환 방법에 대한 영역별 성능비교 결과(단위: pixel).

Table 2. A comparative result of two proposed methods.

	들로네 삼각화 방법	호모그래피 행렬 방법
안쪽영역	3.35	5.14
바깥쪽 영역	6.21	6.05
전체영역	4.76	5.51

(a) The static camera.

(b) The PTZ camera.

그림 13. 실내에서 움직이는 물체 추적 결과 예제 영상.

Fig. 13. An example of indoor object tracking.

(a) The static camera.

(b) The PTZ camera.

그림 14. 실외에서 움직이는 물체 추적 결과 예제 영상.

Fig. 14. An example of outdoor object tracking.

것이다.

그림 13과 그림 14는 제안된 시스템을 통하여 최종적으로 PTZ 카메라에 의하여 제어된 결과를 실내 실험 결과와 실외 실험 결과를 잘 보여주고 있다. 고정된 카메라의 영상에 비하여 PTZ 카메라 영상의 경우 물체의 변별력이 매우 높아짐을 확인할 수 있다.

V. 결론 및 추후 과제

본 논문에서는 기존의 화상 감시 시스템의 문제점을 분석하고, 기존 화상 감시 시스템의 문제점을 개선하면서, 움직이는 물체 추적 알고리즘을 적용하여 좀 더 효과적인 화상 감시 시스템을 제안하였다. 연구를 통해 PTZ 카메라와 고정 카메라를 동시에 활용하여 어떻게 두 카메라를 동기화시키고, 효과적인 화상 감시 시스템에 적용할지에 대하여 기술하였다. 두 카메라의 자동위치 설정은 두 카메라간에 공통된 특징점을 파악하고 이 특징점들을 통해 카메라간 호모그래피 변환 파라미터를 도출하고 이를 상쇄시킬 수 있도록 PTZ 카메라를 조정하여 두 카메라의 시점을 일치시켰다.

고정 카메라에서 추적 알고리즘을 통해 파악한 추적 물체 정보에 대해 실시간으로 PTZ 카메라 추적이 이루어지기 위해서는 고정 카메라상의 좌표를 PTZ 카메라로 변환하고, PTZ 카메라상의 좌표에 대해 적절한 PTZ 카메라 제어 신호로 변환하는 방법을 제안하였다. 두 카메라 좌표 변환 알고리즘은 특징점 쌍을 이용한 호모그래피 행렬 방법과 들로네 삼각화 기법을 활용한 선형 보간법을 활용한 방법을 제안하였고, 실험을 통하여 그 효용성을 입증하였다.

추후 과제로서 고정된 카메라에서 다수의 움직이는 물체가 발생하는 경우 PTZ 카메라 제어를 어떻게 제어할지에 대한 운용전략 등에 대한 연구가 추가적으로 필요하고, PTZ 카메라의 좌표 변환 알고리즘이 좀 제어를 통하여 중심에러가 증가하는 문제를 해결하는 연구가 필요하다.

참고문헌

- [1] R. Horaud, D. Knossow, and M. Michales, "Camera cooperation for achieving visual attention," *Machine Vision and Application*, vol. 16, pp. 331-342, Feb. 2006.
- [2] H. C. Liao and W. Y. Chen, "A dual-PTZ-camera system for visual tracking of a moving target in an open area," *IEEE Advanced Communication Technology (ICACT2009)*, Feb. 2009.
- [3] I. Draganjac, A. Kovacic, D. Ujlaki, and J. Mikulic, "Dual camera surveillance system for control and alarm generation in security applications," *2008 IEEE ISIE (International Symposium on Industrial Electronics)*, Jun. 2008.
- [4] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: a survey," *ACM Comput. Surv.* 38, 4, Article 13, Dec. 2006.
- [5] G. Kim, S. Lee, J. S. Park, G. W. Kim, and J. S. Cho, "Automatic initial position estimation algorithm for dual

camera surveillance systems,” *The Institute of Electronics Engineers of Korea Fall Conference 2011*, pp. 575-578, KAIST, Nov. 2011.

- [6] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, “SURF: Speeded Up Robust Features,” *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359, Apr. 2008.
- [7] D. T. Lee and B. J. Schachter, “Two algorithms for constructing a delaunay triangulation,” *International Journal of Computer and Information Sciences*, vol. 9, no. 3, pp. 219-242, 1980.
- [8] I. Haritaoglu, D. Harwood, and L. Davis, “W4: real-time surveillance of people and their activities,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, *IEEE Transactions on*, vol. 22, no. 8, pp. 809-830, Aug. 2000.
- [9] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Transactions on Pattern Analysis and aMachine Intelligence*, vol. 22, no. 11 pp. 1330-1334, 2000.
- [10] J. H. Kim, T. Zhu, J. Y. Kim, and D. J. Kang, “A method to improve the performance of adaboost algorithm by using mixed weak classifier,” *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 15, no. 5, pp. 457-464, May 2009.
- [11] K. S. Seo and Y. K. Kim, “Scale and rotation robust genetic programming-based corner detectors,” *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 16, no. 4, pp. 339-345, Apr. 2010.

김 기 석

2010년 2월 한국기술교육대학교 인터넷 공학과(공학사). 2010년 3월~현재 한국기술교육대학교 대학원 컴퓨터 공학석사 과정. 관심분야는 Image Processing, Data Association, Programming Language.

이 삭

2012년 3월 한국기술교육대학교 컴퓨터 공학과(공학사). 2012년 3월~현재 한국기술교육대학교 대학원 컴퓨터 공학석사 과정. 관심분야는 Pattern Recognition, Machine Vision System, Multi Dimensional Video Processing.

박 종 섭

2009년 3월~현재 한국기술교육대학교 컴퓨터공학부 학사 과정. 관심분야는 Pattern Recognition, Machine Learning, Artificial Intelligence, Game Programming.

조 재 수

1993년 2월 경북대학교 전자공학과(공학사). 1996년 2월 한국과학기술원 전기 및 전자공과(공학석사). 2001년 2월 한국과학기술원 전기 및 전자공학(공학박사). 2001년 3월~2003년 7월 (주) 삼성전자. 2003년 9월~현재 한국기술교육대학교 컴퓨터공학부 부교수. 관심분야는 Automatic Video Tracking, Visual Surveillance, Pattern Recognition, Machine Learning.