

스테레오 비전을 이용한 선로 장애물 검출에 관한 연구

박기서^o, 이대호^{**}, 박영태^{*}, 김길동^{***}, 오세찬^{***}, 이준희^{****}
경희대학교 전자정보대학^o, 경희대학교 교양학부^{**}, 철도기술연구원^{***}, 건아정보기술^{****}
{arsenide^o, nize, ytpark}@khu.ac.kr, {gdkim, soh}@krri.re.kr, jhlee@keona.co.kr

A Study on Obstacle Detection in Railway using Stereo Vision

Kiseo Park^o, Deaho Lee, Youngtae Park, Gildong Kim, Sehchan Oh, Junhee Lee
Kyung Hee University, Korea Railroad Research Institute, Keon-a Information Technology Co.,Ltd.

요 약

본 논문에서는 스테레오 비전을 이용하여 지하철 선로에 존재하는 장애물을 검출하는 기법을 제안한다. 본 논문에서는 두 대의 아날로그 카메라를 이용하여 좌우 영상을 획득하며, 영역 기반 정합을 이용하여 영상의 불균형을 계산한다. 불균형 영상의 배경을 매 프레임에서 갱신하고 현재 프레임의 불균형과의 차이에 의해 선로상의 장애물을 검출한다. 실내 역사에서 임의로 투척한 신문지, 상자, 인형과 임의로 진입한 사람을 모두 검출하였으며, 한 대의 카메라를 이용하는 기법에서 해결하기 어려운 배경과 비슷한 색상을 가지는 물체도 모두 검출하였다. 따라서 제안하는 기법은 선로상의 위험물을 검출하는 핵심 기술로 이용될 수 있다.

1 서 론

인간의 시각 체계는 서로 다른 위치에서 획득된 두 영상의 차이에 의해 3차원 정보를 인식한다. 이러한 인간의 시각 모델에 기반한 스테레오 비전에 관한 많은 연구가 오랫동안 이루어져 왔다. 이 방법은 초음파와 레이저를 광원으로 사용하여 빛의 비행시간과 속도의 함수로써 거리를 측정하는 것보다 효과적이고 실제 응용 환경의 제약을 적게 받는다는 장점 때문에 의료 영상, 공장 자동화 및 지도 제작에 널리 이용되고 있다[1].

스테레오 비전 기술은 의료현장에서 세심한 로봇팔의 움직임 제어를 위한 기초정보로서 수술대상자의 골격에 대한 인식을 위한 수단과 내시경 촬영 방법으로도 이용할 수 있고, 가상공간을 구현하기 위한 3차원 정보의 제작방법으로도 사용될 수 있으며 군사적으로는 미사일 목표물에 대한 거리 정보 획득에도 유용하게 사용될 수 있다[2,3].

본 논문에서는 2대의 아날로그 카메라로부터 획득된 영상에 대한 스테레오 정합을 이용하여 영상의 불균형을 계산하여 지하철 선로상의 위험 물체를 검출한 기법을 제안한다. 새롭게 진입한 위험 물체를 판별하기 위해서 불균형 영상의 배경을 매 프레임에서 갱신하고, 현재 프레임의 불균형과의 차이를 이용하여 물체를 검출한다. 실내 역사에서 임의로 투척한 신문지, 상자, 인형과 임

의로 진입한 사람을 모두 검출하였으며, 한 대의 카메라를 이용하는 기법에서 해결하기 어려운 배경과 비슷한 색상을 가지는 물체도 모두 검출하였다.

이후 논문의 구성은 다음과 같다. 2장에서는 영역 기반 스테레오 정합에 관련된 기본 이론을 설명하고, 3장에서는 물체 검출 및 위험도 판정을 기법을 기술한다. 또한 4장에서는 실험 결과를 분석하고, 5장에서는 결론과 향후 연구 방향을 제시한다.

2 영역 기반 스테레오 정합

스테레오 정합을 수행하기 위해서는 스테레오 영상이 가지는 고유의 특성을 이용함으로써 좋은 결과를 얻을 수 있는데, 그 제한 사항으로 정합된 값은 서로 유사성을 갖게 되는 유사성의 성질과 변위값이 급격하게 변하지 않게 되는 완만함의 성질, 각 영상의 화소는 동일한 3차원상의 지점을 가리키는 유일성의 성질을 이용한다. 이때 영상의 적절한 특징 추출과 그에 따른 정합 전략, 정합 영역의 크기 및 변이 결정, 경계선 영역 처리 등이 매우 중요하다.

스테레오 영상의 특징은 그림 1과 같이 표현할 수 있으며, 처리의 편리성을 위하여 에피폴라 평면(epipolar plane)을 기준으로 좌우(left, right) 영상의 에피폴라 선(epipolar line)은 동일한 y 축 상에 존재하도록 한다.

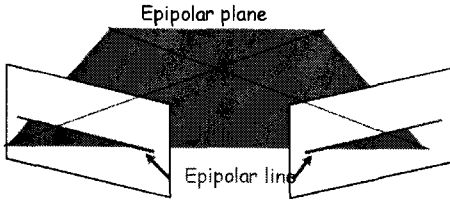


그림 1. 에피폴라 기하

스테레오 정합에 의한 깊이 정보의 산출은 그림 2와 같은 처리 과정을 거친다. 처음의 전처리는 국부 조명 조건의 보상을 위한 과정으로, 본 논문에서는 식(1)과 같은 DC-notch 필터를 이용한다.

$$g(x,y) = f(x,y) - \bar{f}(x,y) \quad (1)$$

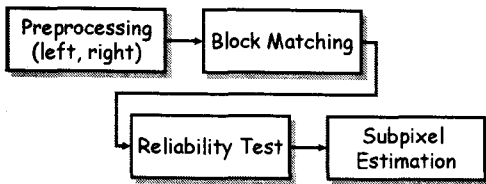


그림 2. 스테레오 정합에 의한 깊이 정보 산출 과정

식(1)에서 $f(x,y)$ 는 원 영상을 의미하고, $\bar{f}(x,y)$ 는 국부 평균을 의미한다. 따라서 필터링후의 영상 $g(x,y)$ 는 국부 평균에 의해 보상된다.

영역 기반 스테레오 정합[4,5]은 기준 영상(왼쪽)의 모든 화소의 위치를 중심으로 하는 블록에 대한 참조 영상(오른쪽)에서 정합도가 높은 블록을 찾는 것이다. 따라서 이러한 블록 정합은 그림 3과 같이 기준 영상에서 정합할 위치의 좌표(x,y)에 대한 블록과 차이가 최소가 되는 블록을 참조 영상에서 찾는다. 그림 3에서, 왼쪽의 (x,y) 좌표와 최소의 차이를 가지는 블록의 중심은 $(x-d,y)$ 가 된다.

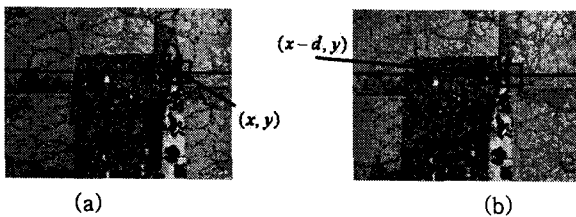


그림 3. 블록 정합
(a) 왼쪽 영상, (b) 오른쪽 영상

이와 같은 블록 정합은 식(2)와 같은 최소 절대 차이(SAD, Sum of Absolute Difference)로 계산할 수 있다. 식(2)에서 $2n+1$ 은 블록의 크기이고, L 은 왼쪽 영상을 의미하고, R 은 오른쪽 영상을 의미한다. 따라서 SAD가 최소가 되는 d 를 정합 위치로 지정할 수 있으며, 이것을 스테레오 영상의 불균형(disparity)이라고 한다.

$$SAD(x,y,d) = \sum_{i,j=-n}^n |L(x+j,y+i) - R(x-d+j,y+i)| \quad (2)$$

식(2)와 같은 SAD에 의한 영상의 불균형의 계산은 모든 화소를 최대의 검색 구간(d_{max})에 대해서 수행해야 한다. 따라서 모든 화소에 대한 식(2)의 계산은 많은 처리 시간을 요구한다. 따라서 신속한 위험 판단을 위하여, 식(2)를 식(3)과 같이 수정하여 계산할 수 있다. 즉, (x,y) 위치의 SAD는 저장된 $(x,y-1)$ 의 SAD로 보다 빠르게 계산할 수 있다.

$$SAD(x,y,d) = SAD(x,y-1,d) + \sum_{j=-n}^n |L(x+j,y+n) - R(x-d+j,y+n)| - \sum_{j=-n}^n |L(x+j,y-n-1) - R(x-d+j,y-n-1)| \quad (3)$$

식(3)의 추가적인 합의 항을 U 에 대한 함수로 표현하면 식(4)와 같이 표현할 수 있으며, U 는 식(5)와 같이 저장된 $(x-1,y)$ 의 U 로 더욱 빠르게 계산할 수 있다.

$$SAD(x,y,d) = SAD(x,y-1,d) + U(x,y,d) \quad (4)$$

$$U(x,y,d) = U(x-1,y,d) + |L(x+n,y+n) - R(x-d+n,y+n)| - |L(x+n,y-n-1) - R(x-d+n,y-n-1)| + |L(x-n-1,y+n) - R(x-d-n-1,y+n)| - |R(x-d-n-1,y-n-1)| \quad (5)$$

영상의 화질이 낮거나 그림 4와 같이 겹쳐진 영역이 발생하는 경우에는 계산된 영상의 불균형은 신뢰하기 어렵다. 그림 4와 같이 겹쳐진 영역은 좌우 영상에서 다르게 나타나므로 실제의 불균형 크기와 일치하지 않는다. 이러한 위치는 SAD의 값이 높은 값으로 나타날 것이다. 따라서 신뢰도를 확인하기 위하여 식(6)과 같이 가장 낮은 불균형 위치(d_{min})와 그 다음으로 낮은 위치(d_i)와의 차이를 이용할 수 있다. 만일 식(6)의 δd 가 낮

으면 d_{min} 의 위치 주위에 비슷한 불균형 위치를 가지므로 신뢰도가 높다고 판정할 수 있다. 또한 식(7)에서 $\delta\epsilon$ 는 d_{min} 에서 계산된 $SAD(\epsilon_{min})$ 주위(d_i)의 SAD의 차이로서, $\delta\epsilon/\epsilon_{min}$ 가 크다면 오류값이 큰 위치로 계산된 것이므로 높은 신뢰도로 인정할 수 있다.

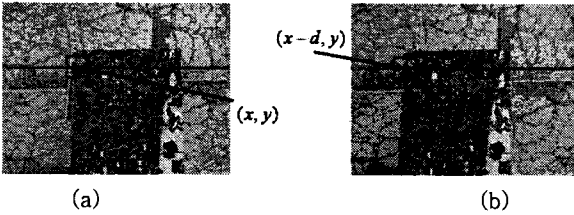


그림 4. 신뢰도가 낮은 위치의 블록
(a) 왼쪽 영상, (b) 오른쪽 영상

$$\delta d = \sum_{i=1}^3 |d_i - d_{min}| \quad (6)$$

$$\delta\epsilon = \sum_{i=1}^3 (\epsilon_i - \epsilon_{min}) \quad (7)$$

식(6)과 식(7)에 의해 신뢰도를 판정한 예는 그림 6과 같다. 그림 5(a)는 그림 4의 영상들에 대해서 불균형을 계산한 것이고, 그림 5(b)는 신뢰도가 낮은 위치를 표시한 것이다. 그림 5(b)에서 검은색으로 표시된 위치는 신뢰도가 낮은 위치로 대부분 겹쳐진 부분에서 발생하는 것을 볼 수 있다.



그림 5. 신뢰도 계산
(a) 불균형 영상, (b) 신뢰도 영상

영상의 불균형의 계산은 정합된 이산적인 위치로 산출된다. 따라서 일반적인 불균형의 크기는 그림 6(a)와 같이 이산적으로 표시된다. 하지만 영상에 대한 불균형 크기는 정합된 위치와 절대적으로 일치하지 않는다. 따라서 불균형은 식(8)과 같이 주위의 불균형 크기에 의해 부화소 추정을 수행하면, 보다 정밀한 불균형 영상을 산출할 수 있다. 그림 6은 이와 같은 부화소 추정에 의

해 보정된 불균형을 보여준다.

$$d_x(x, y) = d_i(x, y) + \frac{f(d_i - 1) - f(d_i + 1)}{2(f(d_i - 1) - 2f(d_i) + f(d_i + 1))} \quad (8)$$



그림 6. 부화소 추정
(a) 불균형 영상, (b) 부화소가 추정된 불균형 영상

3 물체 검출 및 위험도 판정

스테레오 정합만으로 물체를 검출하는 기법은 그림 7과 같이 수행된다. 여기서 물체의 검출은 현재의 불균형 영상과 배경으로 생성된 불균형 영상의 차이를 이용한다. 배경은 아래와 같이 현재 프레임(t)에서의 불균형 영상($d_t(x, y)$)과 배경($B_t(x, y)$)에 가중치(α)의 곱으로 갱신된다.

$$B_t(x, y) = B_{t-1}(x, y) + \alpha(d_t(x, y) - B_t(x, y)) \quad (9)$$

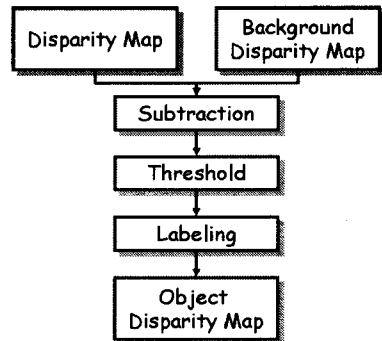


그림 7 불균형 영상에 의한 물체 검출

현재 불균형 영상과 배경의 차이는 식(10)과 같이 계산한다. 이와 같은 배경과의 차이에서 물체의 영역만을 구분하기 위하여 차 영상($D_t(x, y)$)은 임계치에 의해 이진화된다. 이진화된 영상에서 잡은 영역을 포함하고 있으므로, 라벨화(labeling)에 의해 작은 연결 영역을 제거하여 물체의 영역을 최종적으로 검출할 수 있다.

$$D_t(x,y) = |B_t(x,y) - d_t(x,y)| \quad (10)$$

그림 8은 그림 7과 같은 방법에 의해 물체를 검출하는 과정을 보여준다. 그림 8(a)는 왼쪽 영상이며, (b)는 오른쪽 영상이다. 또한 그림 8(c)는 불균형 영상이고, (d)는 부화소 추정에 의해 계산된 불균형 영상이고, (e)는 신뢰도 판정의 결과이며, (f)는 신뢰도가 높은 불균형 영상이다. 그림 8(f)에서 단순히 불균형 영상만으로 물체(뛰어드는 사람)의 영역을 구분하기 어렵다. 이것은 사람의 형상이 복잡한 형상으로 나타나므로 다양한 크기의 블록에 의해서 SAD가 계산되어야만 사람의 형상이 분명히 드러날 수 있다.

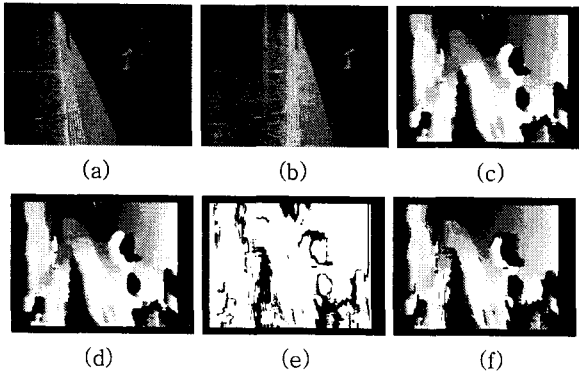


그림 8. 위험 물체 검출 I

- (a) 왼쪽 영상, (b) 오른쪽 영상, (c) 불균형 영상,
- (d) 부화소가 추정된 불균형 영상, (e) 신뢰도 판정,
- (f) 신뢰도가 높은 불균형 영상

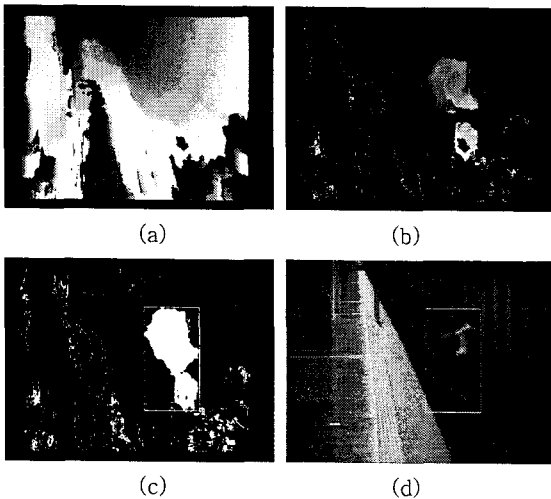


그림 9. 위험 물체 검출 II

- (a) 배경 불균형 영상, (b) 차 영상
- (c) 라벨화에 의한 물체 검출, (d) 검출된 물체

하지만 가변 블록을 적용하지 않더라도, 그림 9과 같이 배경(a)과의 차이에 의해 물체의 형상을 구분할 수 있다. 즉, 배경과 차이가 나는 불균형 영역은 물체가 존재하는 영역이다.

4 실험 결과

실험은 두 개의 아날로그 카메라에서 획득된 영상을 320×240의 회색조 영상으로 처리하였으며, Pentium IV 3.0GHz의 PC에서 $d_{max} = 25$, $n = 25$ 으로 한 프레임을 처리할 때 약 340ms의 시간이 걸린다.

스테레오 정합에 의한 물체의 검출은 그림 10 - 그림 15와 같다. 선로 상에 투척한 신문지, 상자와 인형들을 모두 검출하였으며, 어두운 색의 옷과 밝은 색의 옷을 입은 사람도 모두 검출할 수 있음을 볼 수 있다.

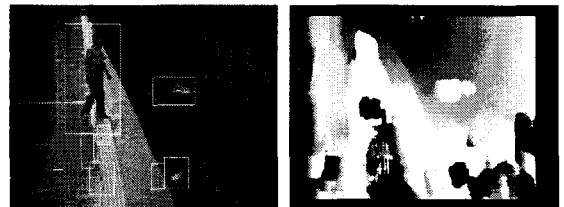


그림 10. 스테레오 정합에 의한 물체 검출(신문지)

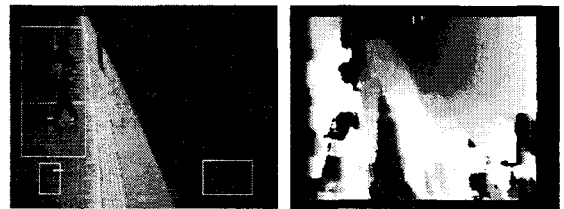


그림 11. 스테레오 정합에 의한 물체 검출(작은 상자)

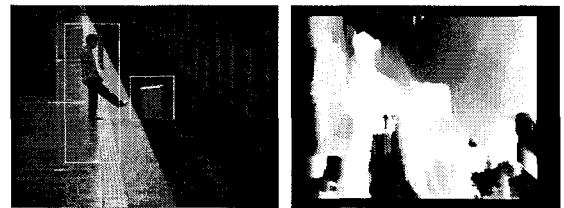


그림 12. 스테레오 정합에 의한 물체 검출(큰 상자)

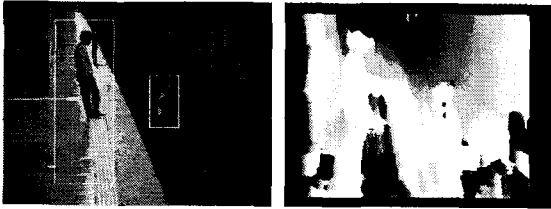


그림 13. 스테레오 정합에 의한 물체 검출(인형)

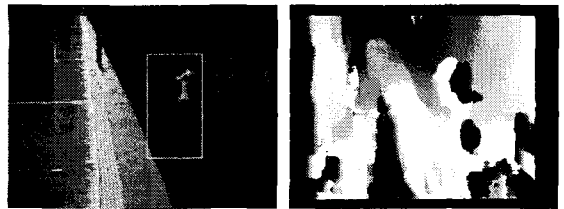


그림 16. 스테레오 정합 결과($d_{max} = 12, n = 12$)

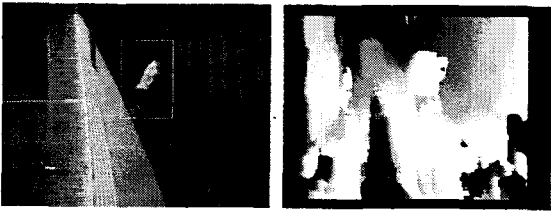


그림 14. 스테레오 정합에 의한 물체 검출
(사람 - 밝은 색상의 옷)

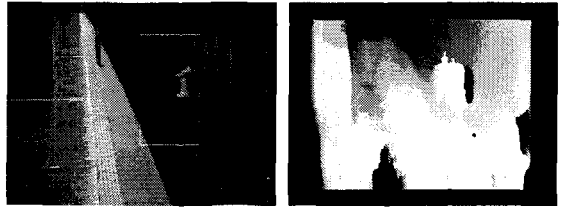


그림 17. 스테레오 정합 결과($d_{max} = 12, n = 20$)

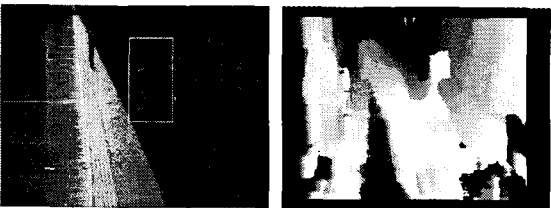


그림 15. 스테레오 정합에 의한 물체 검출
(사람 - 어두운 색상의 옷)

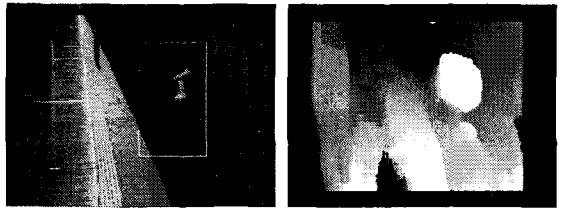


그림 18. 스테레오 정합 결과($d_{max} = 20, n = 20$)

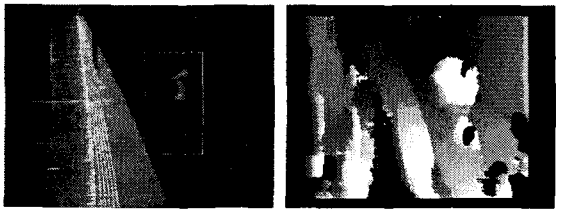


그림 19. 스테레오 정합 결과($d_{max} = 20, n = 12$)

그림 10 - 그림 15의 결과는 스테레오 정합을 위한 식(2)의 d_{max} 와 n 을 각각 12와 12로 설정한 것이다. 여기서 d_{max} 는 정합을 위한 블록을 찾은 범위이며, n 은 블록의 크기이다. 따라서 불균형 크기가 큰 위치에서는 큰 d_{max} 가 필요하다. 또한, 균일한 특성을 가지는 물체 내부에서는 그 크기에 맞는 n 이 사용되면 가장 좋은 결과를 얻을 수 있다. 하지만 처리 시간을 고려한다면 모든 화소에 대해서 동일한 크기의 d_{max} 와 n 이 사용되어야 한다. 이러한 d_{max} 와 n 의 값을 변경한 실험 결과는 그림 16 - 그림 19와 같다. 모두 배경을 이용하여 물체의 영역을 판별할 수 있으나, 물체 영역의 불균형 값은 변화가 있음을 볼 수 있다. 이것은 물체의 형상과 위치에 따라서 가변적인 윈도우에 의한 정합이 정확한 형상을 분리할 수 있음을 보여준다.

5 결론

본 논문에서는 스테레오 정합에 의한 위험 물체의 검출하는 기법을 제안하였으며 실험 결과에서 선로 상에 투척한 신문지, 상자와 인형들을 모두 검출하는 것을 보였다. 또한 어두운 색의 옷과 밝은 색의 옷을 입은 사람도 모두 검출할 수 있음을 볼 수 있었다.

본 논문에서는 스테레오 정합을 이용한 선로상의 위험 물체 검출은 열차의 진입을 고려하지 않고 수행되었다. 하지만 열차의 진입을 고려한다면, 열차의 진입 판단 기법을 같이 사용하여 열차가 진입하지 않은 상황에서만 동작하면 된다. 또한 전체 영역에 대해 모두 불균

형 크기를 측정하지 않고, 프레임 차이나 배경을 이용하여 한 쪽 영상(왼쪽이나 오른쪽)에서 위험 물체를 먼저 식별하고, 식별된 물체 영역에 대해서 불균형 크기를 적용형 윈도우나 가변 윈도우에 의해 판별한다면 매우 우수한 결과를 얻을 수 있을 것이며, 처리 시간도 100ms 이하로 줄일 수 있을 것이다.

따라서 향후 열차 진입을 판별하는 기법과 연결되어 동작시키는 것과 한 쪽 영상에서 검출된 물체 영역에 대한 불균형 크기를 적용형 윈도우나 가변 윈도우로 판별하는 기법이 추가적으로 필요하다. 또한, 측정된 불균형 크기에 의해 물체의 부피를 측정하여 최종적으로 위험도를 판정하는 기법이 필요할 것이다.

"본 연구는 건설교통부의 차세대 첨단 도시철도 시스템 기술개발사업의 지원하에 수행되었습니다."

6 참고문헌

- [1] 임용덕, "스테레오 비전," 한국특허정보원, 2003.
- [2] 이학찬, 박장한, 남궁연, 남궁재찬, "스테레오 영상을 이용한 물체 추적 방법," 대한전자공학회, 전자공학 회논문지-SP, 제39권 제5호, pp. 46 ~ 58, 2002.
- [3] I.Yoda, K.Sakaue "'Ubiquitous Stereo Vision for Controlling Safety on Platforms in Railroad Station,'" IEEJ Tr. on Electronics, Information and Systems, Vol.124, No.3, March, pp.805-811. 2004
- [4] L.D. Stefano, M. Marchionni, and S. Mattocchia, "A fast area-based stereo matching algorithm," Image and Vision Computing, Vol. 22, No. 12, pp. 983-1005, 2004.
- [5] P. Foggia, A. Limongiello and M. Vento, "A real-time stereo-vision system for moving object and obstacle detection in AVG and AMR applications," Proceedings of 7th International Workshop on Computer Architecture for Machine Perception, pp. 58-63, 2005.