

강인한 상관방식 추적기를 이용한 움직이는 물체 추적

A Robust Correlation-based Video Tracking

조재수*, 박동조
(Jae-Soo Cho and Dong-Jo Park)

Abstract : In this paper, a robust correlation-based video tracking is proposed to track a moving object in correlated image sequences. A correlation-based video tracking algorithm seeks to align the incoming target image with the reference target block image, but has critical problems, so called a false-peak problem and a drift phenomenon (correlator walk-off). The false-peak problem is generally caused by highly correlated background pixels with similar intensity of a moving target and the drift phenomenon occurs when tracking errors accumulate from frame to frame because of the nature of the correlation process. At first, the false-peaks problem for the ordinary correlation-based video tracking is investigated using a simple mathematical analysis. And, we will suggest a robust selective-attention correlation measure with a gradient preprocessor combined by a drift removal compensator to overcome the walk-off problem. The drift compensator adaptively controls the template block size according to the target size of interest. The robustness of the proposed method for practical application is demonstrated by simulating two real-image sequences.

Keywords : video tracking, correlation measure, selective-attention

I. 서론

움직이는 물체를 자동으로 추적하는 추적 알고리즘은 레이다나 카메라등의 감지센서 입력 신호를 이용하여 움직이는 물체의 현재 위치, 속도, 물체 크기 등을 실시간으로 추정하는 알고리즘이다. 최근 들어 디지털 비디오 레코더(digital video recorder)가 아날로그 CCTV의 Time-lapse VCR을 대체하면서 빠른 속도로 영상보안시장에 널리 퍼지고 있다. CCTV 산업의 발전과 더불어 보안감시 시스템이 첨단화 되어 가고 있으며, 예전에 군수용 기기에 적용되어온 이러한 움직이는 물체 자동 추적 기능등이 이러한 민수용 감시 장비에도 응용되어, 감시용 장비가 더욱더 지능적이고 첨단화 되어 가고 있다.

비디오 영상을 이용하여 움직이는 물체를 추적하는 방법은 여러 가지 방법이 있지만, 그 중 실시간으로 영상 데이터처리가 가능하여 실제 군수용 장비에 구현되어 온 표적중심 추정기법(centroid estimation of target)[2-5]과 상관방식 추적기법(correlation-based video tracking)[6-10]이 있다. 이 두 방식은 오래전부터 군수용 추적 장비에 널리 응용되어온 대표적인 추적기법이며, 실제로 이러한 두 알고리즘이 다중으로 구현되고 있다. 즉 실제 다양한 응용환경에서 운용이 가능하려면, 상황별로 여러 개의 추적기 중 가장 성능이 좋은 추적방법을 선택할 수 있는 다중모드 추적 알고리즘(multimode tracking algorithm)이 필수적이다.

상관방식 추적기법(correlation tracking method)은 영상에서 목표물의 이동시 발생하는 밝기의 시간적, 공간적 변화량 사이의 관계를 이용하여 움직이는 물체의 이동 정보를

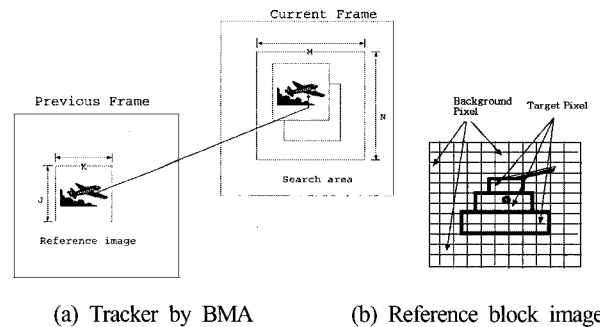


그림 1. 상관방식 추적 개념도.

Fig. 1. Schematic diagram of the correlation.

추출하는 기법으로 시공간 경사법(spatio-temporal method), 블록 정합법(block matching algorithm), 차영상법(difference method)[1]등이 있다.

상관방식 추적 기법 중 가장 대표적인 블록 정합법은 그림 1(a)에서와 같이 이전 영상 프레임의 기준 블록영상을 현재영상의 검색영역으로부터 미소 변위를 구하기 위하여 상관도 함수(correlation function)를 정의하여 이 상관도 함수가 최대가 되는 지점을 이전 프레임으로부터 현재 영상 프레임에 대한 이동 변위로 결정하게 된다. 움직이는 물체 추적뿐만 아니라 영상압축에서 현재 매크로 블록의 움직임을 이전 프레임에서 찾아내는 방법으로 널리 응용되고 있으며, 검색 방법 및 블록정합을 위한 상관도 함수에 대하여 지금까지 많은 연구가 진행되어 왔지만, 대부분이 영상압축을 위한 실시간 구현을 위한 방향으로 연구되어 왔다[6,7,10].

본 논문에서는 먼저 상관방식 추적기에서 널리 사용되어 온 기존의 블록정합 방식에 대한 잘못된 피크(false-peaks) 문제점을 간단한 수학적 해석을 통하여 상관방식 추적기에 대한 문제점을 분석하고, 이러한 문제점을 극복하기 위

* 책임저자(Corresponding Author)

논문접수 : 2004. 2. 25., 채택확정 : 2005. 4. 8.

조재수 : 한국기술교육대학교 인터넷미디어공학부

(jaesoo27@kut.ac.kr)

박동조 : 한국과학기술원 전자전산학부(djpark@ee.kaist.ac.kr)

해 선택적 집중도를 가진 강인한 상관도 함수를 제안하였고, 상관방식 추적기의 또 다른 문제점인 편류현상을 극복하기 위한 방법을 제안하였다. 그리고, 제안된 새로운 상관방식 추적기의 성능을 검증하기 위하여 두개의 테스트 추적영상을 이용하여 그 성능을 검증하였다.

II. 블록정합 상관방식 추적기에 대한 수학적인 해석

움직이는 물체를 포함한 기준 블록영상 $I_r(i, j)$ 는 그림 1(b)에서와 같이 기준 표적영상 $t_r(i, j)$ 과 기준 배경영상 $b_r(i, j)$ 로 구성되어 있으며, 다음식과 같이 수학적으로 모델링 할 수 있다.

$$I_r(i, j) = t_r(i, j) + b_r(i, j) + n(i, j) \quad (1)$$

$$\simeq t_r(i, j) + b_r(i, j) \quad (2)$$

여기서, $n(i, j)$ 는 노이즈를 나타내고, 이러한 노이즈는 무시할 수 있다고 가정한다.

만약 카메라 센서에서 들어오는 센서 입력영상 $I_s(i, j)$ 이 추적하고자 하는 이전의 움직이는 물체를 포함하고 있다고 가정한다면 마찬가지로 현재의 센서 입력영상은 다음식과 같이 표현할 수 있다.

$$I_s(i, j) = t_s(i, j) + b_s(i, j) + n(i, j) \quad (3)$$

$$\simeq t_s(i, j) + b_s(i, j) \quad (4)$$

여기서, $t_s(i, j)$ 는 기준 표적영상($t_r(i, j)$)이 물체의 움직임에 의하여 기하학적으로 변형(geometric distortion)된 표적영상을 나타내고, $b_s(i, j)$ 는 이동물체의 움직임에 의하여 변화된 새로운 배경영상을 의미한다.

(2)와 (4)를 이용하여, 움직이는 물체를 포함한 기준 블록영상과 센서 입력영상의 전체 영상 중 임의의 검색영역에 대한 블록 정합의 상호상관함수(Cross-Correlation Function, CCF)는 (5)와 같이 표현되고, 검색영역 중 (6)과 같이 상호상관함수가 최대가 되는 지점이 현재 입력영상에서 움직이는 물체가 이동한 지점으로 추정할 수 있다.

$$CCF(p, q) = \frac{1}{S_r} \sum_i \sum_j I_r(i, j) \cdot I_s(p + i, q + j) \quad (5)$$

where, $-M_1 \leq p \leq M_1, -M_2 \leq q \leq M_2,$

$$(\hat{p}, \hat{q}) = \arg \max_{(p, q)} CCF(p, q). \quad (6)$$

여기서, S_r 은 기준 블록영상의 전체 사이즈를 나타내고, M_1 과 M_2 는 검색영역을 위해 미리 정해진 정수값이다. (\hat{p}, \hat{q}) 는 상호상관함수에 의하여 추정된 움직이는 물체의 중심위치를 의미한다. (앞으로 편의상 다음과 같이 두개의 합기호를 하나로 사용할 것이다($\sum_i \sum_j = \sum_{i, j}$)).

Definition 1: 기준 표적영상과 현재의 센서 입력영상과의 상호상관함수인 표적-상호상관함수 $CCF_t(p, q)$ 와 기준

배경영상과 현재의 센서 입력영상과의 상호상관함수인 배경-상호상관함수 $CCF_b(p, q)$ 를 다음과 같이 정의한다.

$$CCF_t(p, q) = \frac{1}{S_r} \sum_{i, j} t_r(i, j) \cdot I_s(p + i, q + j) \quad (7)$$

$$CCF_b(p, q) = \frac{1}{S_r} \sum_{i, j} b_r(i, j) \cdot I_s(p + i, q + j) \quad (8)$$

Definition 2: 다음과 같이 기준 표적영상과 센서 입력영상의 상호상관함수인 표적-상호상관함수($CCF_t(p, q)$)가 최대가 되는 지점 (p^*, q^*) 를 정확한 표적 중심이라고 정의한다.

$$(p^*, q^*) = \arg \max_{(p, q)} CCF_t(p, q) \quad (9)$$

정의 1, 2를 통하여, 우리는 다음과 같은 보조정리를 얻을 수 있다.

Lemma 1: 기준 블록영상과 센서 입력영상과의 상호상관함수 $CCF(p, q)$ 는 표적-상호상관함수($CCF_t(p, q)$)와 배경-상호상관함수($CCF_b(p, q)$)의 합으로 표현된다.

$$CCF(p, q) = CCF_t(p, q) + CCF_b(p, q) \quad (10)$$

Proof:

$$CCF(p, q) = \frac{1}{S_r} \sum_{i, j} I_r(i, j) \cdot I_s(p + i, q + j)$$

$$= \frac{1}{S_r} \sum_{i, j} [t_r(i, j) + b_r(i, j)] \cdot I_s(p + i, q + j)$$

$$= \frac{1}{S_r} \left[\sum_{i, j} t_r(i, j) \cdot I_s(p + i, q + j) \right.$$

$$\left. + \sum_{i, j} b_r(i, j) \cdot I_s(p + i, q + j) \right]$$

$$= CCF_t(p, q) + CCF_b(p, q). \quad \blacksquare$$

Remark: 기준 블록영상과 센서 입력영상과의 상호상관함수($CCF(p, q)$)를 이용하여 움직이는 표적중심으로 추정된 지점 (\hat{p}, \hat{q}) 는 정확한 표적중심 (p^*, q^*) 이 아닐 수 있다. 즉 (11)과 같이 잘못된 피크로 인하여 움직이는 표적중심 위치를 잘못 추정할 수 있다는 것이다.

$$\text{If } \arg \max_{(p, q)} CCF(p, q) = \arg \max_{(p, q)} CCF_t(p, q),$$

$$\text{then, } (\hat{p}, \hat{q}) = (p^*, q^*), \quad (11)$$

$$\text{Otherwise, } (\hat{p}, \hat{q}) \neq (p^*, q^*).$$

Lemma 1과 Remark는 기존의 블록정합을 이용한 상관방식 추적기의 문제점인 잘못된 피크 검출로 인한 추적실패의 문제점을 잘 보여주고 있다. 사실, 기준 블록영상의 경우 움직이는 물체의 크기가 계속해서 변하게 되고, 부득이 하게 기준 블록영상 내에 배경영상을 포함하지 않을 수 없는 것이다. 이러한 문제로 인하여, 움직이는 물체를 포함하는 기준 블록의 크기를 최적으로 결정하는 것이 잘못된 피크를 줄이는 하나의 방법이 될 수도 있지만, 최적의 기준 블록영상의

크기를 결정하기는 쉽지가 않다.

블록 정합을 이용하여 움직이는 물체를 좀 더 정확하고, 오랫동안 추적하기 위해서는 앞에서 언급한 배경-상호상관함수의 영향을 없애주는 새로운 상호상관함수가 요구된다.

Definition 3: 가중치 함수 $w(i, j)$ 가 다음과 같은 조건을 만족할 경우, 선택적 집중 가중치(selective-attention weight)라고 정의한다.

$$w(i, j) = \begin{cases} 1, & \text{if } \text{pixel}(i, j) \in \text{Target pixel} \\ 0, & \text{otherwise,} \end{cases} \quad (12)$$

Definition 4: 선택적인 집중도를 가진 상호상관함수 SCCF(Selective-attention Cross-Correlation Function)를 다음과 같이 정의한다.

$$SCCF(p, q) = \frac{1}{S_r} \sum_{i,j} w(i, j) \cdot I_r(i, j) \cdot I_s(p+i, q+j), \quad (13)$$

여기서, $-M_1 \leq p \leq M_1$, $-M_2 \leq q \leq M_2$,

Lemma 2: 선택적인 집중도를 가진 상호상관함수는 표적-상호상관함수 $CCF_t(p, q)$ 와 같다.

$$SCCF(p, q) = CCF_t(p, q). \quad (14)$$

Proof:

$$\begin{aligned} SCCF(p, q) &= \frac{1}{S_r} \sum_{(i,j)} w(i, j) \cdot I_r(i, j) \cdot I_s(p+i, q+j) \\ &= \frac{1}{S_r} [\sum_{(i,j)} w(i, j) \cdot t_r(i, j) \cdot I_s(p+i, q+j) + \\ &\quad \sum_{(i,j)} w(i, j) \cdot b_r(i, j) \cdot I_s(p+i, q+j)] \\ &= \frac{1}{S_r} \sum_{(i,j)} t_r(i, j) \cdot I_s(p+i, q+j) = CCF_t(p, q). \end{aligned}$$

Theorem: 선택적인 집중도를 가진 상호상관함수를 이용하여 추정된 움직임을 물체 중심 예측 지점 (\hat{p}, \hat{q}) 은 정확한 표적중심 (p^*, q^*) 와 일치한다.

$$(\hat{p}, \hat{q}) = (p^*, q^*). \quad (15)$$

Proof: Lemma 2에 의하여,

$$\begin{aligned} (\hat{p}, \hat{q}) &= \arg \max_{(p,q)} SCCF(p, q) \\ &= \arg \max_{(p,q)} CCF_t(p, q) \\ &= (p^*, q^*). \end{aligned}$$

이상으로 기존의블록정합에 의한 상관방식 추적기는 기준 블록영상내의 배경영상의 영향으로 잘못된 표적 중심을 추정할 확률이 매우 높음을 알 수 있다. 이러한 문제점을 극복하기 위해 3장에서는 선택적인 집중도를 가진 강인한 상관함수를 제안하였다.

III. 선택적인 집중도를 가진 강인한 상관함수

2장에서 기존의 블록정합에 의한 움직이는 표적 추적에 배경-상호상관함수에 의하여 잘못된 위치를 검출할 확률이 매우 높음을 간단한 수학적 해석을 통하여 알 수 있었다.

그리고, 선택적 집중 가중치를 이용한 선택적 집중 상호상관함수를 통하여 이러한 문제점을 해결할 수 있었다. 그러나, 실제적으로 정합을 하는 각 픽셀이 표적 픽셀인지 배경 픽셀인지를 정확하게 알 수 없다.

이에 본 장에서는 선택적 집중 상호상관함수의 성질을 잘 이용하면서 선택적 집중 가중치의 결정에 대한 문제점을 해결할 수 있는 강인한 선택적 집중 상호상관함수를 다음과 같이 제안한다.

$$SCCF(p, q) = \frac{1}{S_r} \sum_{i,j} \text{AND}[\hat{I}_r(i, j), \hat{I}_s(p+i, q+j)] \quad (16)$$

$$(\hat{p}, \hat{q}) = \arg \max_{(p,q)} SCCF(p, q) \quad (17)$$

여기서, $\hat{I}_r(i, j)$ 는 경계선 추출을 위한 미분 전처리(gradient preprocessing)과정을 거친 이진 기준 블록영상을 의미하고, $\hat{I}_s(i, j)$ 는 센서 입력영상을 마찬가지로 경계선 추출을 위한 미분 전처리 과정을 거친 이진 영상을 의미한다. 즉 식 (16)은 기존의 상호상관함수가 검색영역과 기준 블록영상간의 모든 픽셀에 대하여 동일한 가중치를 가지고 상호상관함수를 계산하였지만, 본 논문에서 제안하는 강인한 선택적 집중 상호상관함수는 기준 블록영상의 경계선에 해당하는 픽셀과 검색영상의 경계선 픽셀간의 AND 연산을 통하여, 경계선 부분에 해당하는 픽셀만 상관함수 계산에 포함함으로써 선택적 집중도를 구현한 것이다. 그림 2는 제안된 선택적 집중도를 가진 상관함수를 이용한 강인한 상관방식 추적기에 대한 블록도를 나타내었다. 인덱스 n 은 n 번째 영상프레임을 나타내는 시간변수이다.

센서 입력영상과 기준 블록영상의 전처리는 기준 블록영상안에 존재하는 배경부분의 상관함수 기여도를 줄이고, 실제 목표물의 경계부분만으로 정합하기 위해서 미분 전처리(gradient preprocessing)를 사용하였다. 영상에서 경계선을 검출하는 방법은 여러 가지가 있지만 여기서는 (18)과 같은 미분 연산자를 사용하였다[11].

$$|\nabla| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad (18)$$

기준 블록영상의 갱신은 (19)를 이용하여 이전의 기준 블록영상에서 새로운 기준 블록영상과의 재귀적 필터를 이용하여 생성하게 된다. 이 과정은 그림 3에 나타내었다.

$$I_r(i, j, n) = \mu \cdot G(i, j, n) + (1 - \mu) \cdot I_r(i, j, n-1) \quad (19)$$

여기서, μ 는 기준 블록영상의 시정수에 해당하는 재귀적 필터계수이다.

일반적으로 움직이는 물체의 움직임이 빨라서 움직이는 물체의 모습이 빨리 변하는 경우는 시정수의 값을 크게 한다. 그리고, $G(i, j, n)$ 은 입력영상에서 선택적 집중 상호상관함수에 의하여 추적된 표적 중심에서 이전의 기준영상과 같은 크기로 추출된 블록영상을 나타낸다. 그림 3의 편류보상은 다음 4장에 자세하게 설명하였다.

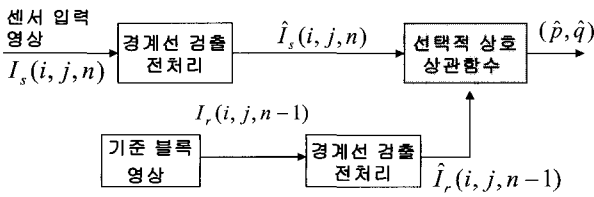


그림 2. 제안된 선택적 집중도 상관함수 추적기.
Fig. 2. Proposed selective-attention cross-correlation tracker.

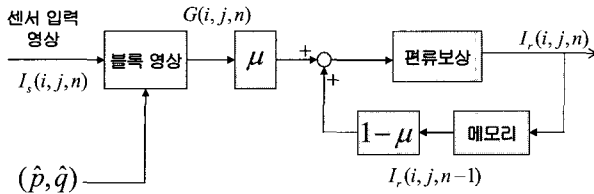


그림 3. 기준 블록영상 갱신에 관한 블록도.
Fig. 3. Block diagram of the reference block image update.

IV. 기준 블록영상의 적응적인 크기 조절 및 편류보상

블록정합에 의한 상관방식 추적기의 문제점 중 또 다른 하나는 Aim point로 사용되는 기준 블록영상의 중심이 추적 이 진행되면서 표적이 기준 영상에서 조금씩 벗어나게 된다. 이 현상을 변위 편류현상(drift phenomenon)이라고 표현 한다.

기준 블록영상에 대한 크기 조절과 편류 보상을 위해 기 준 블록영상을 경계선 추출 처리 과정을 거친 이진영상 ($\hat{I}_r(i, j)$)에 대하여, 행 방향과 열 방향으로 프로젝트션 히스 토그램(projection histogram)을 이용하여 이러한 문제점을 해 결할 수 있다.

$$H(n, r) = \lambda \cdot H(n-1, r) + (1-\lambda) \cdot h(n, r) \quad (20)$$

$$H(n, c) = \lambda \cdot H(n-1, c) + (1-\lambda) \cdot h(n, c)$$

여기서, $h(n, r)$ 과 $h(n, c)$ 는 n 프레임 기준 블록영상의 경계선(에지) 이진영상에 대한 행과 열 방향으로 프로젝트션 된 히스토그램을 나타내고, $H(n, r)$ 과 $H(n, c)$ 는 시정수 λ 를 이용한 필터링된 프로젝트션 히스토그램을 의미한다. 그 리고, r 과 c 는 각각 $0 \leq r \leq GR$, $0 \leq c \leq GC$ 의 값을 가지며, GR 과 GC 는 각각 기준 블록영상의 행과 열의 크기를 의미한다. 필터링된 프로젝트션 히스토그램을 이용하 여 기준 영상내의 표적 크기와 표적의 중심을 계산할 수 있다. 필터링된 히스토그램에서 다음과 같은 값을 얻을 수 있다.

- r_{max}, r_{min} : 기준 블록영상내에서 행 방향으로 표적이 위 치한 부분 중 가장 큰 행 값과 가장 작은 행 값을 의미한다.
- c_{max}, c_{min} : 기준 블록 영상내에서 열 방향으로 표적이 위 치한 부분 중 가장 큰 열 값과 가장 작은 열 값을 나타낸다.

기준 블록영상을 경계선 처리후 이진 영상에서 프로젝트션 히스토그램을 추출하고, 이전의 프로젝트션 히스토그램과 필 터링한 후, 열방향과 행방향으로 물체가 위치한 최대 최소

값을 얻는다면 움직이는 물체의 중심값과 물체 크기를 다음 과 같이 계산할 수 있다.

$$RSize(n) = r_{max}(n) - r_{min}(n) \quad (21)$$

$$CSize(n) = c_{max}(n) - c_{min}(n) \quad (22)$$

$$RCenter(n) = \frac{r_{max}(n) + r_{min}(n)}{2} \quad (23)$$

$$CCenter(n) = \frac{c_{max}(n) + c_{min}(n)}{2} \quad (24)$$

여기서, $RSize(n)$ 과 $CSize(n)$ 은 각각 행 방향, 열 방 향으로 움직이는 표적의 크기를 의미하고, $RCenter(n)$ 과 $CCenter(n)$ 은 각각 기준 블록영상에서의 움직이는 물체의 중심값을 나타낸다. 최종적으로 위에서 정의된 값을 이용하여 기준 블록영상의 행 방향 크기 및 열 방향 크기 를 다음과 같이 결정할 수 있다.

$$RRS(n) = RSize(n) + \sigma_r(n) \quad (25)$$

$$RCS(n) = CSize(n) + \sigma_c(n) \quad (26)$$

$RRS(n)$ 과 $RCS(n)$ 은 각각 영상 n 프레임에서 기준 블록영상의 행 방향, 열 방향 크기를 의미하며, $\sigma_r(n)$ 과 $\sigma_c(n)$ 은 행 방향과 열 방향으로 배경을 포함할 수도 있는 허용오차를 의미한다. 이 허용오차가 크면 클수록 기준 블 록영상 내에 배경영상을 많이 포함하게 되고, 허용오차가 작으면 기준 블록영상 내에 배경영상을 적게 포함할 수 있 겠지만, 움직이는 물체의 주위 일부분을 기준 블록영상내에 서 포함하지 못할 수도 있어, 궁극적으로 쉽게 추적에 실패 할 수 있다. 그래서, 일반적으로 움직이는 물체를 모두 포 함하고, 주위의 배경 영상을 기준 블록영상내에서 포함하지 않을 수 없는 것이다. 본 논문에서는 이 허용오차를 기준 블록영상의 행방향 및 열 방향 크기의 5~10% 정도로 하여 실험하였다.

매 프레임마다 위에서 정의한 값들을 계산하여 기준 블 록영상의 크기 조절 및 편류 보상을 한다. 즉 블록 정합이 끝나고 새로운 기준 블록영상을 생성한 후, 새로운 기준 블 록영상에 대한 경계선 검출을 하고, 프로젝트션 히스토그램을 계산한후, 위에서 정의한 값들을 계산한 후 움직이는 물체 를 기준 블록영상의 중심에 두고 크기도 위에서 정의한 값 으로 새로운 기준 블록영상을 생성하게 된다. 기준 블록영 상에 대한 필터링된 프로젝트션 히스토그램을 이용하는 이유 는 좀더 안정적이고, 주위의 배경 영향을 줄이기 위함이다. 기준 블록영상에 대한 크기 조절은 움직이는 물체의 속도 에 한계가 있기 때문에 물체의 크기도 그 속도에 상응하는 크기 변화를 반영하였다. 이전 프레임들에서의 움직이는 물 체 크기의 평균값과 현재 프레임에서의 움직이는 물체크기 계산값 사이의 허용오차안에 있을 경우만 기준 블록영상의 크기를 조절하였다.

V. 실험결과 및 고찰

여러 가지 실제 영상을 이용하여 컴퓨터 시뮬레이션하였 으며, 기존의 다른 상관함수를 이용한 상관방식 추적기와

제안된 방법과의 성능을 비교검증하였다. 먼저 실험에 사용된 테스트 영상들을 그림 4에 나타내었다. 그림 4(a) 영상은 카메라가 고정된 CCD 테스트 영상으로 움직이는 장난감 자동차 영상이며, 그림 4(b) 영상은 FLIR 영상으로 카메라도 움직이는 탱크 영상으로, 두 영상 모두 640×480 크기의 영상을 15 frame/sec의 속도로 획득한 영상이다. CCD 영상은 총 90 프레임, FLIR 영상은 총 178 프레임이다.

움직이는 물체 추적은 움직이는 물체 감지 및 획득 부분과 추적(움직이는 물체 정보 추출), 그리고 궁극적으로 카메라 제어 부분으로 구성되어 있다. 본 논문에서는 움직이는 물체 감지 및 획득 방법은 논문 [12]에서 제안된 디스터번스 맵과 무빙 윈도우 개념을 이용하여, 초기 움직이는 물체를 감지하고, 획득하였다. (더 자세한 내용은 논문 [12]를 참조바랍니다.)

움직이는 표적을 추적하기 위해서 다음과 같은 등속도 운동 상태방정식과 측정방정식을 이용하여 추적 필터인 칼만필터를 이용하여 상태값을 예측하였다.

$$x(n+1) = \begin{pmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{pmatrix} x(n) + \begin{pmatrix} \frac{T^2}{2} & 0 \\ T & 0 \\ 0 & \frac{T^2}{2} \\ 0 & T \end{pmatrix} v(n) \quad (27)$$

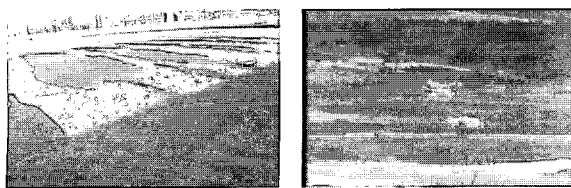
$$z(n) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} x(n) + w(n) \quad (28)$$

(27)은 상태방정식이고, (28)은 측정방정식이며, 상태값은 움직이는 물체의 위치와 속도로 이루어지고, T값은 샘플링 시간을 의미하며, $v(n)$ 은 화이트 가우시안 노이즈이고, $w(n)$ 은 측정 노이즈를 의미한다.

본 논문에서 움직이는 물체 추적을 위해 제안한 선택적 집중 상호상관함수의 성능을 검증하기 위해, 기존의 블록정합 알고리즘에서 사용되어온 여러 가지 상관도 함수들과 비교실험을 하였다. 여러 가지 상관도 함수[6-8]는 CCF (Cross-Correlation Function), MAE(Mean-Absolute Error), NCCF (Normalized Cross-Correlation Function), NMAE(Normalized Mean Absolute Error), MSE(Mean Squared Error), BBME (Binary Block Matching with Edge map)와 비교하였다. 기존의 상관함수들에 대한 수식은 다음과 같다.

$$MAE(p, q) = \sum_{(i,j)} |I_r(i, j) - I_s(p+i, q+j)| \quad (29)$$

$$(\hat{p}, \hat{q}) = arg \min_{(p,q)} MAE(p, q) \quad (30)$$



(a) CCD car image

(b) FLIR tank image

그림 4. 실험 영상.

Fig. 4. Test image sequences.

$$NMAE(p, q) = \sum_{(i,j)} |I_r(i, j) - \bar{I}_r - I_s(i+p, j+q) + \bar{I}_s(p, q)| \quad (31)$$

$$(\hat{p}, \hat{q}) = arg \min_{(p,q)} NMAE(p, q) \quad (32)$$

$$\bar{I}_r = \frac{1}{NM} \sum_{k=1}^N \sum_{l=1}^M I_r(k, l), \bar{I}_s(p, q) = \frac{1}{NM} \sum_{k=1}^N \sum_{l=1}^M I_s(p+k, q+l) \quad (33)$$

$$NCCF(p, q) = \frac{[\sum_{i,j} I_r(i, j) \cdot I_s(i+p, j+q)]^2}{[\sum_{i,j} I_r^2(i, j)][\sum_{i,j} I_s^2(i+p, j+q)]} \quad (34)$$

$$(\hat{p}, \hat{q}) = arg \max_{(p,q)} NCCF(p, q) \quad (35)$$

$$MSE(p, q) = \frac{1}{S_r} \sum_{(i,j)} [I_r(i, j) - I_s(i+p, j+q)]^2 \quad (36)$$

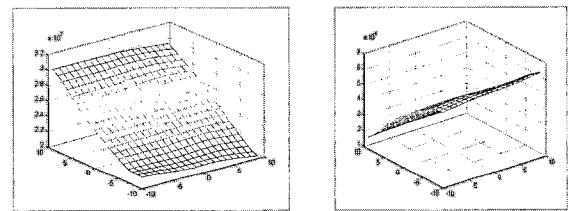
$$(\hat{p}, \hat{q}) = arg \min_{(p,q)} MSE(p, q) \quad (37)$$

$$BBME(p, q) = \sum_{(i,j)} XOR[\hat{I}_r(i, j), \hat{I}_s(i+p, j+q)] \quad (38)$$

$$(\hat{p}, \hat{q}) = arg \min_{(p,q)} BBME(p, q) \quad (39)$$

상관함수 BBME의 경우 본 논문에서 제안한 경계선 검출을 위한 이전 영상을 이용한 상관함수이다.

2장에서 수학적 해석을 통하여 지적하였듯이, 배경이 복잡할 경우 상호상관함수에 대한 전역최대점(global-maximum point)이 실제 목표물이 존재하는 부분이 아닌 다른 부분에서 나타날 수 있다는 것이다. 실제로 이러한 전역 최대점에 대한 문제를 확인하기 위하여 그림 5에서 실제 자동차 CCD 영상 (그림 4(a))을 이용하여 상호상관함수 CCF와 상관함수 MAE에 대한 분포도를 나타내었다. 그림 5(a)는 이전 프레임에서 목표물이 포함된 기준 블록영상을 이용하



(a) CCF surface

(b) MAE surface

그림 5. 상호상관함수와 MAE 분포도.

Fig. 5. Cross-correlation surface and MAE surface.

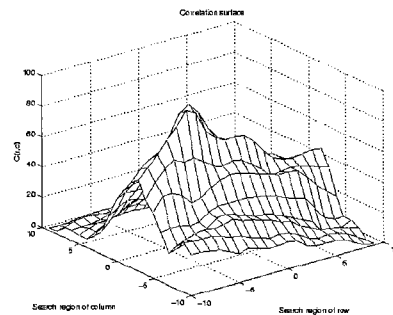


그림 6. 제안된 선택적 집중 상호상관함수 분포도.

Fig. 6. Proposed selective-attention cross-correlation surface.

여 현재 프레임에서의 검색 영역 20×20에 대한 상호상관함수 CCF에 대한 분포도를 나타내고, 그림 5(b)는 상관함수 MAE에 대한 검색영역에서의 분포도를 나타내었다. 이 결과에서 알 수 있듯이 기존의 블록정합법에서 가장 많이 사용되는 상호상관함수 CCF나 상관함수 MAE를 그대로 사용할 경우, 배경과 움직이는 표적의 밝기값에 대한 상관성이 매우 높을 경우 추적이 거의 불가능함을 알 수 있다. 이에 반하여, 본 논문에서 제안한 선택적 집중 상호상관함수를 이용한 분포도를 그림 6에 나타내었으며, 그림 6에서 보는 바와 같이 전역 최대점(global maximum point)가 뚜렷하게 존재함을 알 수 있다.

추적 결과에 대한 추적 성능 비교를 위하여 다음과 같은 두 가지 추적 성능 값을 정의하였다.

$$Error(n) = \sqrt{(i(n) - \hat{i}(n|n-1))^2 + (j(n) - \hat{j}(n|n-1))^2} \quad (40)$$

$$RMSE = \frac{1}{n} \sum_{k=1}^n Error(k) \quad (41)$$

여기서 $Error(n)$ 는 n 프레임에서 실제 표적의 중심위치와 예측된 표적의 중심위치와의 추적오차거리를 나타내는 값이고, $RMSE$ 는 n 프레임까지의 추적 오차평균값을 의미한다. $(i(n), j(n))$ 은 n 프레임에서의 실제 표적의 중심좌표를 의미하고, $(\hat{i}(n|n-1), \hat{j}(n|n-1))$ 은 추적 필터에 의해 추정된 표적의 중심위치를 나타내는 값이다.

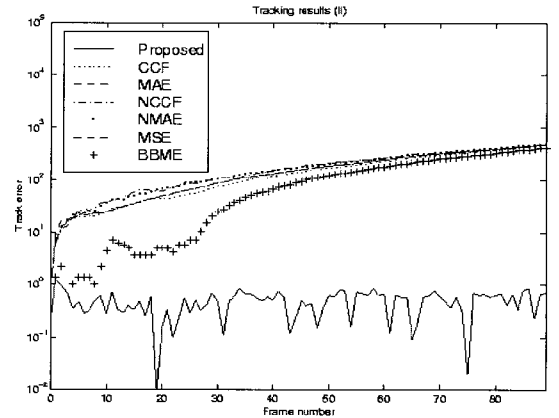
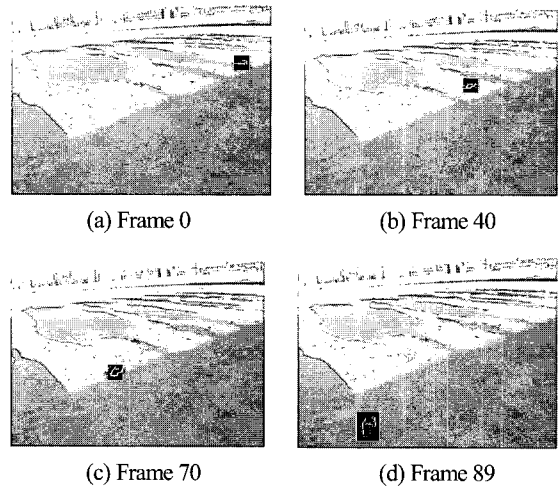
제안된 방법으로 CCD 자동차 영상에 대한 추적결과를 그림 7에 나타내었다. 그림 7(a),(b),(c),(d)는 각각 추적 결과 영상을 보여주고, (e)는 다른 여러 상관함수와 제안된 상관함수를 이용하여 추적한 전체 추적 에러 결과를 보여준다. 그림 7(f)는 매 프레임마다 움직이는 물체의 크기 변화에 적응적으로 변화하고 있는 기준 블록영상에 대한 열과 행방향의 크기를 나타내었다. 마찬가지로 그림 8은 FLIR 탱크 영상에 대한 추적결과를 나타내었다. CCD 자동차 영상과 FLIR 탱크영상에 대한 $RMSE$ 값은 표 1에 나타내었다.

필터링에 사용된 시정수는 0.5를 사용하였으며, 실험한 실제 추적영상은 기존의 블록정합법으로는 몇 프레임에서 추적을 실패하였지만 제안된 방법으로는 추적이 잘 되었으며, 기준 블록영상에 대한 크기도 적응적으로 조절됨을 알 수 있다. 그림 9는 기준 블록영상에 존재하는 노이즈를 제거하기 위해 재귀적 필터링방법을 사용하였는데, 필터링을 사용한 경우와 사용하지 않은 경우에 대한 차이를 잘 보여주고 있다. 그리고, 그림 10은 기준 블록영상에서 편류보상을 위해 행방향과 열방향으로의 프로젝션 히스토그램을 보여주고 있다.

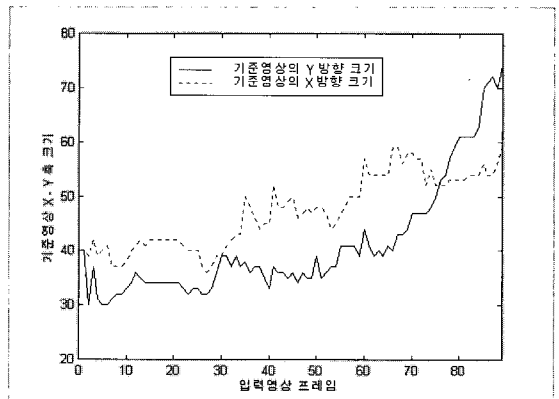
실험결과에서 알 수 있듯이 추적의 특성상 추적 중 한 프레임이라도 잘못된 피크를 검출할 경우 추적에 실패하게 된다. 두 실험 영상뿐 만 아니라 다른 여러 실험 영상에서도 본 논문에서 제안된 선택적 상관도 함수(즉 움직이는 표적의 경계선 부분만을 이용한 에지 정합)와 상관방식 추적 알고리즘의 편류현상을 개선한 방법이 추적 성능이 매우 좋음을 확인할 수 있었다.

VI. 결론

본 논문에서는 움직이는 물체 추적 알고리즘 중 기존의 블록정합 알고리즘을 이용한 상관방식 추적기의 문제점을 간단한 수학적 모델링을 통하여 잘못된 피크 검출 문제점을 분석하였고, 이러한 잘못된 피크 검출 문제점을 극복하기 위해, 선택적 집중도를 가진 강인한 상관함수를 제안하였다.

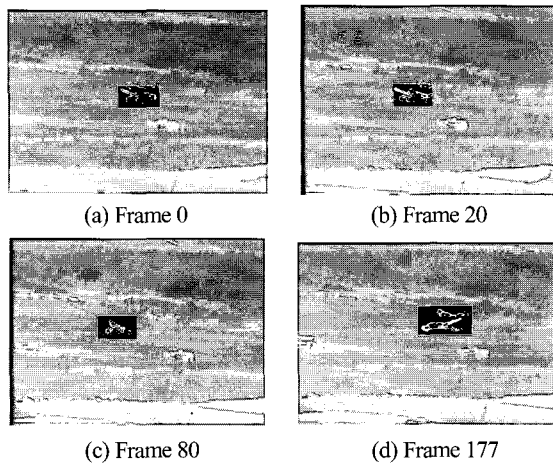


(e) Overall tracking results



(f) Size change of the reference block image

그림 7. CCD 자동차 영상 추적 결과.
Fig. 7. Tracking results of CCD car image.

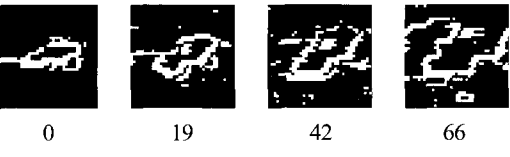
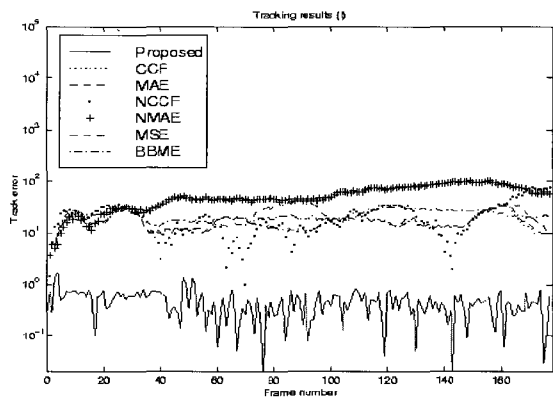


(a) Frame 0

(b) Frame 20

(c) Frame 80

(d) Frame 177



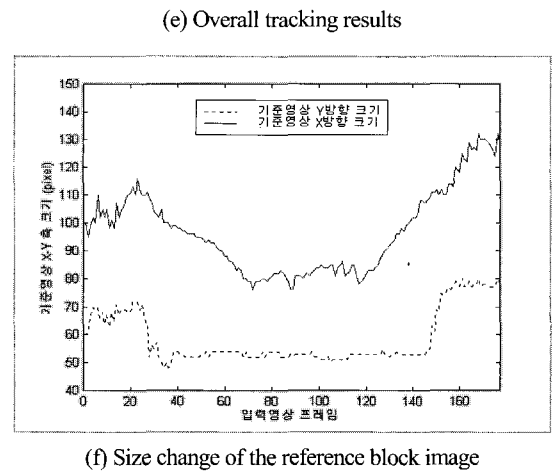
(a) Reference block image without recursive filtering



(b) Reference block image with recursive filtering

그림 9. 기준 블록영상 생성에 관한 비교 결과.

Fig. 9. Reference block image generation.



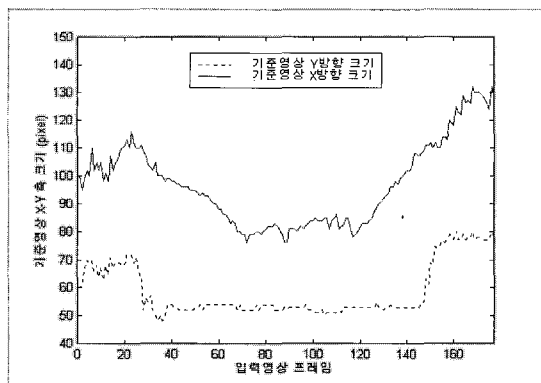
(a) Reference block image

(b) Projection histogram

그림 10. 기준 블록영상에 대한 프로젝션 히스토그램.

Fig. 10. Projection histogram of a reference image.

(e) Overall tracking results



(f) Size change of the reference block image

그림 8. FLIR 탱크 영상 추적 결과.

Fig. 8. Tracking results of FLIR tank image.

표 1. 평균 추적 오차값.

Table 1. RMSE value.

상관함수/실제영상	자동차 영상	탱크 영상
CCF	172.4	16.82
MAE	183.68	16.90
NCCF	196.80	16.10
NMAE	200.09	34.62
MSE	184.01	19.93
BBME	128.87	17.73
제안된 방법	0.50	0.49

제안된 선택적 집중도를 가진 상호상관함수가 기존의 여러 상관도 함수보다 잘못된 피크가 발생할 확률이 매우 적게 되는 이유는 다음과 같다. 전처리 후의 에지는 세가지 원인에 의한 값이다. 움직이는 표적의 에지 성분, 움직이지 않는 배경물체의 에지 부분, 그리고 잡음(noise)에 의한 에지 등 3가지중 하나이다. 잡음에 의한 에지 성분은 연속된 두 영상의 같은 화소 부분에 연속적으로 나타날 확률이 매우 적다(연속된 두 영상의 잡음은 서로 독립적이다). 그리고 기준 블록영상의 움직이지 않는 배경 부분의 에지 성분은 그 다음 프레임에서 움직이는 표적의 움직임으로 인하여 기준 블록영상의 배경부분의 에지와 검색 영상에서의 움직이지 않는 배경 부분의 에지가 정합되지 않는다. 덧붙여서, 기준 블록영상의 갱신 과정에서 이전의 기준 블록영상과의 필터링에 의하여 잡음성분에 의한 에지와 배경 부분에 의한 에지는 그림 9에서 보는 것과 같이 기준 블록영상 내에서는 거의 없어지게 된다. 이러한 이유로 기준 블록영상의 에지와 검색영상의 에지가 정합되는 부분은 움직이는 표적에 의한 에지성분만 결과적으로 상관도에 영향을 미치는 것이고, 본 논문에서 제안하는 선택적 집중도를 가진 상관함수가 되는 것이다.

블록정합에 의한 상관방식 추적기의 또 다른 문제점인 편류현상을 극복하기 위한 방법으로 적응적으로 기준블록영상의 크기와 기준 블록영상을 갱신함으로써 편류현상을 극복하는 방법을 제안하였다. 이렇게 제안된 방법은 기존의 고정된 CCD 카메라 영상과 움직이는 FLIR 카메라 실험 영상을 이용하여 그 성능을 검증하였다.

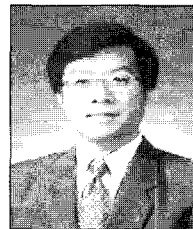
참고문헌

- [1] W. L. Martin and J. K. Aggarwal, "Extraction of moving object descriptions via differencing," *Computer Graphics and Image Processing*, vol. 18, pp. 188-201, 1982.
- [2] E. Oron, A. Kumar, and Y. Bar-Shalom, "Precision tracking with segmentation for imaging sensors," *IEEE Trans. on Aerospace Electron. Syst.*, vol. 29, no. 3, pp. 977-986, July 1993.
- [3] A. Kumar, Y. Bar-Shalom, and E. Oron, "Precision tracking based on segmentation with optimal layering for imaging sensors," *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 17, no. 2, pp. 182-188, Feb. 1995.
- [4] J. S. Cho, D. J. Kim, and D. J. Park, "Robust centroid target tracker based on novel distance features in cluttered image sequences," *IEICE Trans. Inf. and Syst.* vol. E83-D, no. 12, pp. 2142-2151, Dec. 2000.
- [5] B. G. Kim and D. J. Park, "Novel precision target detection with adaptive thresholding for dynamic image segmentation," *Machine Vision and Applications*, vol. 1, no. 12, P. 259 - 270, Mar. 2001.
- [6] H. Gharavi and M. Mills, "Block matching motion estimation algorithm-new results," *IEEE Trans. Circuit. Syst.*, vol. 37, no. 5, pp. 649-651, May 1990.
- [7] M. J. Chen, L. G. Chen, T. D. Chiueh, and Y. P. Lee, "A new block-matching criterion for motion estimation and its implementation," *IEEE Trans. Cir. Syst. Video Technol.*, vol. 5, no. 3, pp. 231-236, June 1995.
- [8] D. A. Montera, S. K. Rogers, D. W. Ruck and M. E. Oxley, "Object tracking through adaptive correlation," *Optical Engineering*, vol. 33, no. 1, pp. 294-302, Jan. 1994.
- [9] A. M. Peacock, S. Matsunaga, D. Renshaw, J. Hannah and A. Murray, "Reference block updating when tracking with block matching algorithm," *Electronics Letters*, vol. 36, no. 4, pp. 309-310, Feb. 2000.
- [10] K. Nickels and S. Hutchinson, "Estimation uncertainty in SSD-based feature tracking," *Image and Vision Computing*, vol. 20, no. 1, pp. 47-58, 2002.
- [11] E. Gose, R. Johnsonbaugh, *Pattern Recognition and Image Analysis*, Prentice Hall, 1996.
- [12] 조재수, 추길환, "복잡한 영상신호에서 디스터번스 맵을 이용한 움직이는 물체 자동감지, 획득 및 추적," 대한전자공학회 2003 정보 및 제어학술회의 논문집, pp. 199-202, Nov. 2003.



조재수

1993년 경북대학교 전자공학과 졸업(학사). 1996년 한국과학기술원(KAIST) 전기및전자공학과 졸업(공학석사). 2001년 한국과학기술원(KAIST) 전자전산학부 전기및전자공학 졸업(공학박사). 2001년~2003년 (주)삼성전자 DVS 사업부 책임연구원. 2003년~현재 한국기술교육대학교 인터넷미디어공학부 전임강사.



박동조

1976년 서울대학교 전자공학과 졸업(학사). 1981년 UCLA(University of California, Los Angeles) 전자공학 졸업(공학석사). 1984년 UCLA(University of California, Los Angeles) 전자공학 졸업(공학박사). 1984년~1985년 한국전자통신연구소(ETRI) 선임연구원. 1986년~현재 한국과학기술원(KAIST) 전자전산학부 교수.