Journal of The Korea Society of Computer and Information
Vol. 21 No. 11, pp. 49-57, November 2016

www.ksci.re.kr
http://dx.doi.org/10.9708/jksci.2016.21.11.049

Visual Tracking using Weighted Discriminative Correlation Filter

Tae-Eun Song*, Kyung-Hyun Jang**

Abstract

In this paper, we propose the novel tracking method which uses the weighted discriminative correlation filter (DCF). We also propose the PSPR instead of conventional PSR as tracker performance evaluation method. The proposed tracking method uses multiple DCF to estimates the target position. In addition, our proposed method reflects more weights on the correlation response of the tracker which is expected to have more performance using PSPR. While existing multi-DCF-based tracker calculates the final correlation response by directly summing correlation responses from each tracker, the proposed method acquires the final correlation response by weighted combining of correlation responses from the selected trackers robust to given environment. Accordingly, the proposed method can provide high performance tracking in various and complex background compared to multi-DCF based tracker. Through a series of tracking experiments for various video data, the presented method showed better performance than a single feature-based tracker and also than a multi-DCF based tracker.

▸ Keyword : target tracking, PSR, discriminative correlation filter

I. Introduction

영상기반 표적 추적 기술은 자동제어시스템, 감시정찰 및 보안 등 다양한 분야에서 활용되는 필수적인 기술이다. 이러한 표적 추적 분야에서 단조로운 배경의 표적을 추적하는 것에 비해, 복잡한 배경의 표적을 추적하는 것은 매우 도전적이다[1-3]. 일반적으로 복잡한 배경에서 표적을 추적할 때는 표적 및 배경의 특징을 정확하게 예측할 수 없기 때문에, 복잡한 배경에서 강인한 추적을 위하여 환경 및 표적의 변화에 강인한 특징을 활용한 표적 추적 기법들이 제안되어왔다.

최근에 제시된 DCF(discriminative correlation filter) 기반 추적 방법들은 복잡한 배경에서 표적을 추적할 때 뛰어난 성능을 보여주고 있다[4]. DCF기반 추적기와 전통적인 상관추적기법의 가장 큰 차이점은 상관필터를 업데이트하는 방법에 있다. 일반적인 상관추적기법은 입력된 영상에서 추출한 특징을 상관필터학습에 바로 사용하는 반면, DCF기반 추적기는 비용함수를 최소화시키도록 상관필터를 갱신한다. 이렇게 업데이트된 상관필터는 주어진 추적 환경 및 표적의 변화에 강인하게 학습되어 복잡한 배경에서도 높은 추적성능을 보여준다. 이러한 DCF기반 추적기는 대표적으로 ASEF(Average of Synthetic Exact Filters)[5], MOSSE(Minimum Output Sum of Squared Error)[6], KCF(Kernelized Correlation Filter)[7] 등이 있다.

하지만 이와 같은 상관필터 학습방법에도 불구하고 한 특징에서 학습된 상관필터는 일부 환경 변화 및 표적 변화에만 강인하며, 다양한 환경 변화에 모두 강인한 상관필터를 생성하는 것은 극히 어려운 일이다.

이러한 한계점을 극복하기 위해서 단일 특징을 이용한 DCF 기반 추적방법은 다차원 특징을 적용한 방법으로 확장되었다[8-10]. 다차원 특징으로 확장된 multi-DCF 추적 방법은 상호보완이 되는 몇 가지 특징 및 상관필터를 이용해서 다양한 환경 변화에 강인한 추적기를 만든다. multi-DCF 방법은 각 특징들의 상관응답을 구한 후, 모든 상관응답을 누적하여 최종 상관응답을 생성한다. 이러한 방법을 통하여, 일부 특징에서 잘못된 표적위치를 계산하더라도 나머지 다수의 상관응답이 표적의 위치를 올바르게

• First Author: Tae-Eun Song, Corresponding Author: Kyung-Hyun Jang
*Tae-Eun Song (te09.song@hanwha.com), Image Sensor Team, Hanwha Systems Co., Ltd.
**Kyung-Hyun Jang (kyunghyun.jang@hanwha.com), Image Sensor Team, Hanwha Systems Co., Ltd.
• Received: 2016. 10. 28, Revised: 2016. 11. 08, Accepted: 2016. 11. 25.

추정할 수 있다. 반면에, 각 차원의 상관응답들을 조합하는 과정이 단순 누적 합이기 때문에 상관응답의 값이 큰 특징에 의하여 결과가 많은 영향을 받는다. 또한 서로 다른 특징으로 계산된 상관응답 값의 크기를 정규화과정 없이 적용하는 것은 바람직하지 못한 결과를 초래할 수 있다. 따라서 상관응답을 조합하는 과정에 각 상관응답에 대한 정규화 과정이 필요하다.

본 논문은 서로 다른 특징을 이용한 DCF기반 추적기 여러 개를 조합하는 새로운 방법을 제안한다. 제안하는 방법은 먼저 영상으로부터 서로 다른 특징들을 추출한다. 이 때 사용되는 특징은 조밀한 맵을 추출할 수 있으면 어떠한 특징추출방법도 적용 가능하다. 추출된 각 특징을 이용하여 상관추적을 수행하고 획득된 상관응답들에 대하여 제안하는 정량적 평가방법으로 신뢰도를 평가한다. 이렇게 얻어진 신뢰도 값을 기반으로 다중 상관응답들을 가중치 융합한다. 제안하는 기법은 주어진 환경에 강인한 추적기와 그렇지 않은 추적기를 선별하고, 추적성능이 상대적으로 더 뛰어날 것이라고 예상되는 추적기를 추적결과에 더 많이 가중치를 반영한다. 이렇게 함으로써 모든 상관응답을 단순 합하는 기존 방법에 비하여 좀 더 다양하고 복잡한 배경에서도 추적성능을 향상시킬 수 있다. 본 논문에서는 단일 특징을 이용한 대표적인 DCF기반 추적기인 MOSSE 추적기[6]를 제안하는 방법의 기본 추적기로 사용하고, 제안한 새로운 조합방법으로 다차원 가중합 DCF(multi-WDCF)기반 추적기를 구성하였다.

다음 절에서는 먼저 기본으로 사용하는 단일 특징 기반 추적기인 MOSSE 추적기와 이를 다차원으로 확장한 버전 multi-DCF 추적기에 대해서 설명하고, 기존에 주로 사용해 오던 추적기 평가방법인 PSR(Peak to Sidelobe Ratio)에 대해 설명한다. 다음으로 제안하는 추적기의 새로운 신뢰도 평가 방법인 PSPR(Peak to Sidelobe Peak Ratio)과 이를 이용한 추적기 조합방법을 설명한다. 마지막으로 제안한 기법에 대하여 다양한 샘플 영상에서 기존 MOSSE 및 multi-DCF기반 추적기와 추적 성능 비교 결과를 제시한다.

II. Preliminaries

1. Related works

1.1 MOSSE(Minimum Output Sum of Squared Error)

상관추적에서 추적기는 새로운 프레임에서 표적의 위치를 검색하기 위해 상관필터를 갱신시킨다. 갱신 과정에서 시간 t 동안에 표적이 포함된 부분영상 $f_1,\dots,f_t$ 이 학습샘플로 사용된다.

각 학습샘플이 목적으로 하는 상관 응답 결과를 $g_1,\dots,g_t$ 라고 한다면, 시간 t 에서 목적으로 하는 상관 응답 결과 g_t를 얻을 수 있는 최적의 상관필터를 계산하기 위하여 비용함수 식(1)을 사용한다. 제곱오차합을 의미하는 식(1)을 최소화하면 원하는 상관필터를 얻을 수 있다.

$$err = \sum_{i=1}^{t} \| h_t \odot f_i - g_i \|^2 = \frac{1}{MN} \sum_{i=1}^{t} \| \overline{H}_t F_i - G_i \|^2 \quad (1)$$

여기에서 $\odot$ 는 상관관계(correlation) 연산자이고, 함수 h_t, f_i, g_i 는 모두 $M \times N$ 크기이다. 대문자는 각 해당 함수의 이산푸리에변환(DFT)을 나타낸다. $\overline{H}_t$는 복소수 공액 연산을 의미한다. 식(1)을 최소화하는 H_t는 다음과 같다.

$$H_t = \frac{\sum_{i=1}^{t} \overline{G}_i F_i}{\sum_{i=1}^{t} \overline{F}_i F_i} \quad (2)$$

목적으로 하는 상관 응답 결과 g_i 는 가우시안 함수를 이용하여 극대점 위치가 f_i 에 포함된 표적의 중심위치와 일치하도록 생성한다. 실제 구현에서는 새로운 영상 f_t 가 주어지면, H_t 는 지난 영상에서 획득된 H_{t-1} 와 가중평균의 형태로 갱신하게 된다.

새로운 영상 z 가 주어졌을 때, 상관응답 cr 은 식(3)과 같이 계산된다.

$$cr = DFT^{-1}\{\overline{H}_t Z\} \quad (3)$$

여기서 DFT^{-1} 는 역 이산 푸리에변환 연산이다.

새로운 표적의 위치는 상관응답 cr 의 최대값의 위치로 추정한다. 학습 및 표적위치를 찾는 과정에서 고속 푸리에 변환(FFT)를 도입하여 효과적으로 처리시간 성능을 높일 수 있다.

1.2 Multi-dimensional DCF

위에서 설명한 MOSSE 추적기는 최근 다차원 특징에 적용될 수 있도록 확장되었다[8-10]. 이러한 다차원 확장버전 추적기에는 영상밝기, HOG 등 조밀한 표현이 가능한 어떠한 특징도 적용될 수 있다.

d-차원의 특징으로 표현된 영상이 주어졌을 때, f 를 전체 영상에서 추출한 특징 중에서 표적이 포함된 일부영역이라고 하고, f^l 은 f 에서 l 번째 차원만 추출한 영상으로 가정하면 ($l \in \{1,\dots,d\}$), 다음의 비용함수 식(4)를 최소화함으로써 최적의 상관필터인 h를 구할 수 있다.

$$err = \left\| \sum_{l=1}^{d} h^l \odot f^l - g \right\|^2 + \lambda \sum_{l=1}^{d} \| h^l \|^2 \quad (4)$$

여기에서 λ는 정규화 항의 효과를 조절하는 변수이고, g 는 f 에 포함된 표적의 위치와 연관되어 생성된 목적으로 하는 상

관 응답 결과이다.

식 (4)의 비용함수를 최소화하는 상관필터 H는 다음과 같이 얻을 수 있다.

$$H^l = \frac{\overline{G}F^l}{\sum_{i=1}^{d}\overline{F}_i F_i + \lambda} \tag{5}$$

이렇게 획득된 H는 주어진 하나의 학습샘플에 대해 계산된 상관필터이다. 실제 구현에서 연속해서 입력되는 학습샘플에 대해서 최적의 상관필터를 근사화하기 위해, MOSSE 와 마찬가지로 상관필터를 가중평균 형태로 갱신한다.

새로운 영상 z 가 주어지면 상관응답을 다음 식(6)과 같이 계산한다.

$$cr = DFT^{-1}\left\{\frac{\sum_{l=1}^{d}\overline{GF^l}Z^l}{\sum_{l=1}^{d}\overline{F^l}F^l + \lambda}\right\} \tag{6}$$

여기서 상관응답 cr 의 최대값의 위치가 새로운 표적의 이동 위치가 된다. Fig. 1은 multi-DCF기반 추적기가 다중 특징 및 다중 상관필터를 이용하여 상관응답을 구하는 과정을 계략적으로 보여준다.

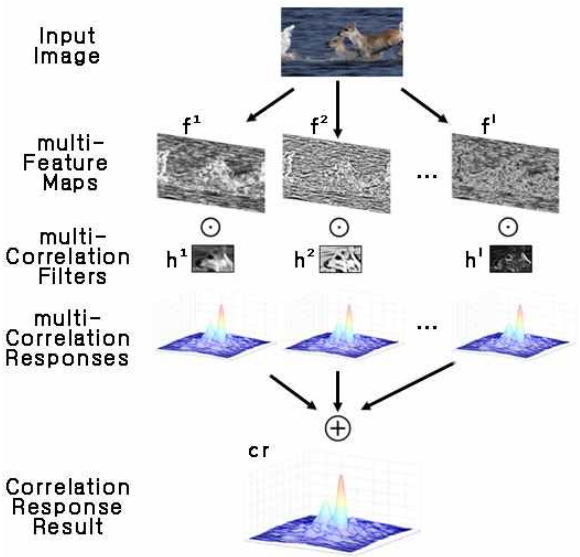

Fig. 1. Multi-DCF Tracking Method using multi-features

multi-DCF는 최종 상관응답 결과를 얻기 위하여 식 (6)과 같이 서로 다른 특징 및 상관필터로 계산된 상관응답들을 누적한다. 그러나 각 차원의 상관응답들은 서로 다른 특징으로부터 계산되었기 때문에 정규화 과정 없이 상관응답들을 누적하게 되면 바람직하지 못한 누적 결과를 야기할 수 있다.

1.3 PSR(Peak to Sidelobe Ratio)

상관추적 기법에서는 주어진 상관응답을 평가하기 위한 수치로써 PSR를 주로 사용해 왔다[6]. PSR값이 높을수록 추적결과의 신뢰도가 높다는 것을 의미한다. 따라서 추적의 성공/실패를 판단하는데 이 값이 사용되기도 한다.

주어진 상관응답으로부터 PSR을 계산하기 위해서 Fig. 2와 같이 최대상관응답을 가지는 극대점($\max_{peak}$)과 그 주변 영역을 제외한 나머지부분(sidelobe)으로 분리하고, PSR을 다음 식 (7)과 같이 정의한다.

$$PSR = \frac{\max_{peak} - mean_{sidelobe}}{\sigma_{sidelobe}} \tag{7}$$

여기에서 $\max_{peak}$는 상관응답의 최대값, $mean_{sidelobe}$와 $\sigma_{sidelobe}$는 sidelobe 영역의 평균과 표준편차를 나타낸다.

PSR의 값이 커지면 배경의 평균 유사도에 비해 표적의 유사도가 상대적으로 더 높아진다. 즉 전반적으로 배경의 유사도에 비해 표적의 유사도가 커지기 때문에 추적결과의 신뢰도가 높아진다. 반면 배경의 전반적인 유사도를 의미하는 평균 유사도를 이용하기 때문에 특정 지점에서 높은 유사도가 나타나도 신뢰도 결과에 별다른 영향을 주지 않는다는 단점이 있다.

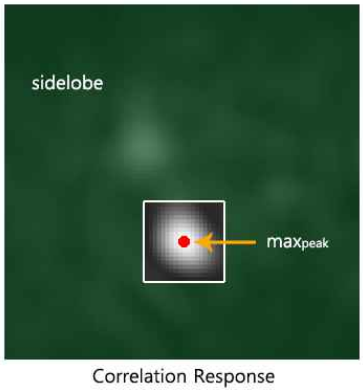

Fig. 2. $\max_{peak}$ and sidelobe

실제 추적기에 적용할 때는 PSR을 결정하기 위해 극대점 주변 제외 영역의 크기를 11×11로 고정시키기도 하고[6], 추적물체의 크기를 추정하는 추적기에서는 제외 영역의 크기를 가변하도록 응용하기도 한다[10].

III. The Proposed Scheme

1. PSPR(Peak to Sidelobe Peak Ratio)

실제 표적과 비슷한 모양의 유사표적이 배경에 존재할 경우, 유사표적이 상관응답에 주는 영향은 유사표적의 위치와 그 주

변영역에 제한된다. 따라서 상관응답의 신뢰도를 정량적으로 평가하기 위해서 상관응답 전체영역의 분포특성을 참조하기보다는 지역적 분포특성을 참조하여야 한다. 또한 추적기가 표적의 위치를 결정할 때, 상관응답의 최대값을 참조한다. 이 때문에 유사표적이 위치한 영역에서 획득한 상관응답의 최대값과 실제 표적이 위치한 영역에서 획득한 상관응답의 최대값 사이의 차이가 추적 성공/실패 판단의 중요한 척도가 될 수 있다.

본 논문에서는 이러한 특성을 고려하여 상관응답을 평가하기 위한 수치로써 기존에 주로 쓰이는 PSR과 다른 새로운 방법인 PSPR을 제안한다. 제안하는 SPR은 추적결과의 신뢰도를 평가하기 위해 여러 상관응답의 통계적 특성을 의미하는 $mean_{sidelobe}$이나 $\sigma_{sidelobe}$를 이용하지 않고, 지역적 분포 특성인 sidelobe 영역 최대값과 전체영역의 최대값을 이용하여 계산할 수 있다.

주어진 상관응답으로부터 PSPR을 계산하기 위해 PSR과 마찬가지로 최대 상관응답을 가지는 극대점(peak)과 그 주변 영역을 제외한 나머지부분(sidelobe)으로 분리하고, PSPR을 다음 식(8)과 같이 정의한다.

$$PSPR = \frac{\max_{peak}}{\max_{sidelobe}} \tag{8}$$

여기에서 $\max_{peak}$는 상관응답의 최대값이고 $\max_{sidelobe}$는 $\max_{peak}$주변 영역을 제외한 나머지부분의 최대값을 의미한다. Fig. 3은 기존 PSR과 제안하는 PSPR에 대한 차이를 나타낸다.

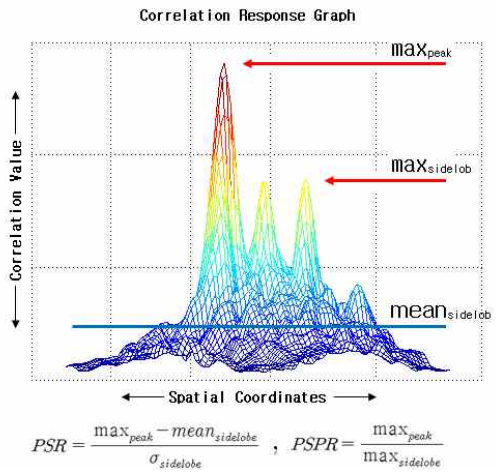

Fig. 3. PSR vs PSPR

2. Probability-transformation and Combination of Correlation Responses

기존의 다차원 확장 상관추적기 multi-DCF는 각 차원의 상관응답들을 모두 같은 가중치를 적용하여 합한 결과를 최종 상관응답으로 취한다. 하지만 각 특징에 해당하는 상관응답은 추출한 특징의 특성에 따라 그 값의 범위가 상당히 차이를 보인다. 이 때문에 추적결과는 상대적으로 높은 상관응답을 가지는 특징에 더 많은 영향을 받는다. 반대로 상대적으로 낮은 상관응답을 가지는 특징들은 추적 결과에 적은 기여를 한다. 이와 같은 현상은 특징의 특성 때문에 발생한 결과로 추적 결과의 신뢰도와는 관련이 없다. 따라서, 최종 추적결과를 도출할 때 추적 성능이 상대적으로 뛰어난 추적기의 결과가 무시될 가능성이 있다.

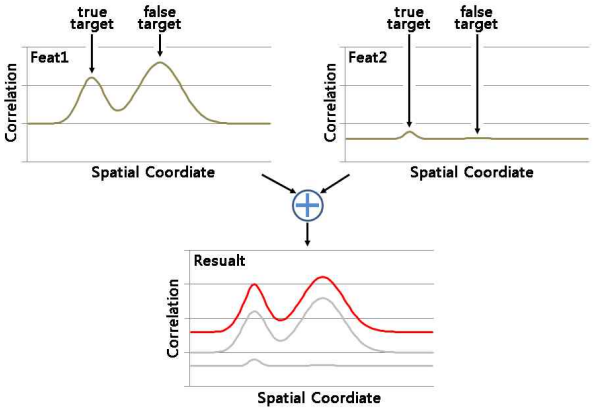

Fig. 4. Combination of correlation responses on multi-DCF

앞서 설명한 상황의 한 예를 Fig. 4에 2개의 상관응답 프로파일로 간략히 나타내었다. 특징1의 상관응답 그래프를 보면 2개의 극대점을 볼 수 있다. 여기서 실제 표적은 왼쪽 극대점에 위치해 있다고 하자. 하지만 특징1 상관응답의 최대점의 위치는 오른쪽 극대점이므로 잘못된 결과를 나타내고 있다. 특징2의 상관응답 그래프는 실제 표적의 위치에 상관응답 최대값을 갖고 있지만, 특징1의 상관응답에 비해 상대적으로 낮은 값을 가지고 있다. 이러한 경우 두 개의 상관응답을 누적하더라도 결과는 잘못된 위치에 극대점을 갖는다. 이는 특징1이 특징2의 상관응답 보다 상대적으로 높은 값에 분포해 있기 때문이다.

이와 같은 문제를 해결하기 위해서 각 차원의 상관응답 값의 범위를 정규화하고, 추적기의 성능에 따라 각 차원의 상관응답을 가중합을 하여 최종 상관응답을 조합하는 방법을 제안한다. 상관응답은 학습된 표적이 해당 위치에 있을 가능성의 척도로 해석할 수 있다. 따라서 상관응답을 전체 합이 1이 되도록 정규화하여, 확률분포 형태로 표현한다. 이때, 상관응답은 음수값을 가질 수 있으므로 먼저, 절대값을 취한다.

$$p^l_{m,n} = \frac{abs(cr^l_{m,n})}{\sum_{i=1}^{M}\sum_{j=1}^{N}(abs(cr^l_{i,j}))} \tag{9}$$

여기서 M,N은 영상의 픽셀크기이고, m,n은 픽셀좌표를 의미한다.

위의 식(9)으로 구해진 각 차원의 확률분포는 신뢰도에 따라 다음과 식(10)과 같이 조합하여 최종 확률분포를 생성한다.

$$p_{m,n} = \sum_{l=1}^{d} \pi^l p^l_{m,n} \tag{10}$$

여기서 π^l는 $0 \le \pi^l \le 1$범위를 가지고, $\sum_{l=1}^{d}\pi^l = 1$ 조건을 만족하는 값으로, 각 차원 추적기의 신뢰도를 의미하는 값이다. 앞서 제안한 PSPR을 이용하여 다음 식(11)과 같이 정의하면, 위 조건을 만족하는 π^l 를 구할 수 있다.

$$\pi^l = \frac{(pspr^l)^\alpha}{\sum_{i=1}^{d}(pspr^i)^\alpha} \tag{11}$$

여기서 $pspr^l$은 l 차원의 확률분포에서 구한 PSPR값이고, $\alpha(\le 0)$ 는 각 차원의 확률분포 p^l이 전체 확률분포 p에 반영되는 정도를 조절하기 위한 변수이다. α가 높은 값을 가질수록 높은 PSPR을 가진 확률분포가 더 많은 기여도를 가진다. 반대로 α가 0값을 가지면 모든 확률분포가 같은 가중치로 더해지게 된다. Fig. 5에 앞서 설명한 과정을 간략화 하였다. 먼저 서로 다른 특징1, 2로 계산된 상관응답들을 확률분포 형태로 정규화 한다. 이 과정에서 상관응답의 레벨 및 변화량은 서로 비교 가능한 수치로 변환된다. 다음으로 가중치 π^l를 곱하여 신뢰도가 높은 특징2의 상관응답을 강조하고 신뢰도가 낮은 특징1의 상관응답은 감쇄시킨다. 결과 확률분포를 보면 multi-DCF가 사용한 단순 누적한 상관응답 결과와 다르게 실제 표적을 나타내는 왼쪽 극대점이 최대상관응답 값을 가지고 있다.

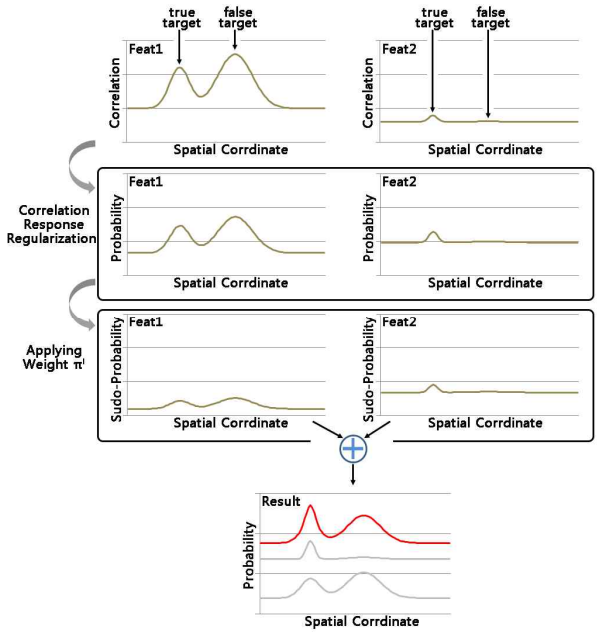

Fig. 5. Combination of correlation responses on multi-WDCF

3. Low Performance Tracker Elimination Method

π^l 이 낮은 값을 가지면 표적과 배경의 구분을 분명히 하지 못하는 추적기일 가능성이 높다. 이러한 추적기는 최종 확률분포를 구할 때, 제외하여 추적성능에 영향을 미치지 못하도록 한다. 이를 위해서 π^l 를 오름차순으로 정렬하고 낮은 순서대로 π^l 의 누적합을 구한다. 누적합이 $\beta(0<\beta<1)$ 보다 작을 때까지의 π^l 는 그 값을 0으로 변경한다.

4. Brief Outline of Proposed Algorithm

제안한 multi-WDCF기반 추적기법의 알고리즘 수행 과정을 요약하면 다음과 같다.

제안 알고리즘 수행 과정

1) 입력 : t 번째 영상 Img_t, 이전 표적 위치 pos_{t-1}
   이전 상관필터 H_{t-1}
2) 영상 Img_t 로부터 d-차원 특징 $f_t = \{f_t^1, \dots, f_t^d\}$ 추출
3) f_t와 H_{t-1}를 이용하여 상관응답 $cr_t^{1,\dots,d}$ 계산
4) $cr_t^{1,\dots,d}$의 정규화 및 $pspr_t^{1,\dots,d}$, 가중치 $\pi_t^{1,\dots,d}$ 계산
5) 최종 상관응답 cr_t 생성
6) cr_t이 최대값을 가지는 위치로 pos_t 설정
7) pos_t 에 극대점을 갖는 목적 상관응답 g_t생성
8) g_t, f_t, H_{t-1} 를 이용하여 상관필터 H_t 생성 및 갱신
9) 출력 : 추정 표적 위치 pos_t, 상관필터 H_t

제안하는 알고리즘의 전체 흐름도를 Fig. 6에 나타내었다.

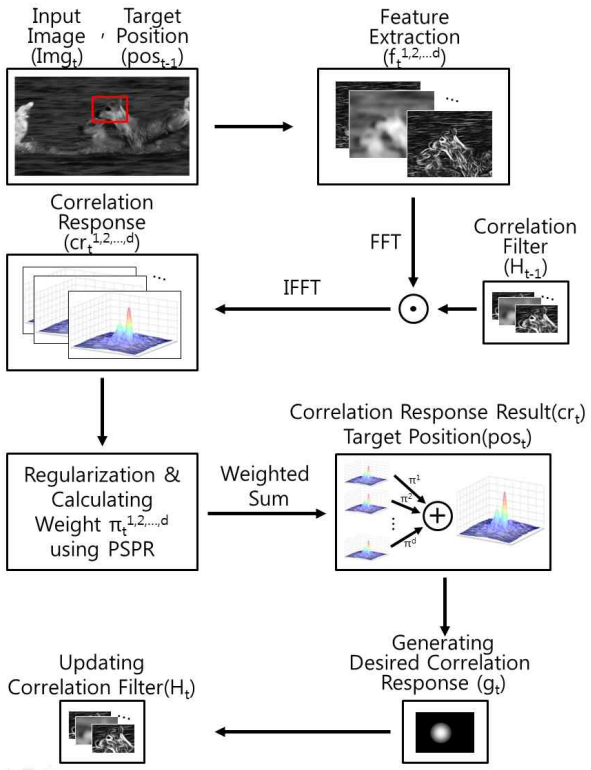

Fig. 6. Flowchart of proposed tracking method

IV. Experiment

1. Evaluation Methodology

표적추적 성능측정 수치로 중심위치오차(Center Position Error), 거리정밀도(Distance Precision), 겹침정밀도(Overlap Precision)을 사용하였다.

중심위치오차는 식 (12)와 같이 추정된 표적 위치 cp_T 와 실제표적위치 cp_G와의 유클리디안 거리를 평균한 값이다.

$$CPE = \frac{1}{N}\sum_{i=1}^{N}\triangle\left(cp_T^i, cp_G^i\right) \tag{12}$$

여기서 $\triangle$는 유클리디안 거리 연산자이고, N은 전체 프레임의 개수이다.

거리정밀도는 식 (13)과 같이 전체 프레임 중에서 임계값보다 작은 유클리디안 거리 오차를 가지는 프레임이 차지하는 비율값이다. 이 값은 추적의 성공률을 의미하는 지표이다.

$$DP = \frac{1}{N}\sum_{i=1}^{N}\triangle_{th}\left(cp_T^i, cp_G^i\right), \tag{13}$$
$$\triangle_{th}(cp_T, cp_G) = \begin{cases} 1, & \text{if}\ \triangle(cp_T, cp_G) \le th,\ th \in R \\ 0, & otherwise \end{cases}$$

겹침정밀도는 식 (14)와 같이 실표적영역 A_G과 추정된 표적영역 A_T의 합집합 대비 교집합 영역의 비율이 임계값보다 높은 프레임이 전체 프레임 중에서 차지하는 비율을 나타낸다. 이 값은 추적의 정확도를 의미하는 지표이다.

$$OP = \frac{1}{N}\sum_{i=1}^{N}\Theta_{th}\left(A_T^i, A_G^i\right), \tag{14}$$
$$\Theta_{th}(A_T, A_G) = \begin{cases} 1, & \text{if}\ \dfrac{A_T \cap A_G}{A_T \cup A_G} \le th,\ th \in R \\ 0, & otherwise \end{cases}$$

2. Experimental Setup

2.1 Test environment

본 논문에서 제시한 추적 방식은 Matlab R2013a 을 이용하여 구현하였다. 그리고 모든 실험은 Intel Xeon CPU E31270 @ 3.40 GHz, 8.00GB Memory, Windows 7 64bit 조건에서 수행하였다.

2.2 Datasets

실험에 사용된 영상은 최근 추적기 평가에 사용된 영상 데이터세트[11]를 사용하였다. 해당 추적 평가용 데이터세트는 영상의 특성에 따라 다양한 태그를 붙여놓았다. 그 중에서 대상표적이 화면 밖으로 완전히 사라짐을 나타내는 OV태그가 있는 영상을 제외하였다. 또한 본 논문에서 제시한 multi-WDCF 추적기는 크기변화를 고려하지 않기 때문에 표적의 크기변화가 심함을 나타내는 SV 태그가 붙은 영상 또한 제외하여 총 18개의 영상을 실험에 사용하였다. Fig. 7는 성능평가에 사용된 영상의 첫 번째 프레임과 초기 표적의 위치를 나타낸다.

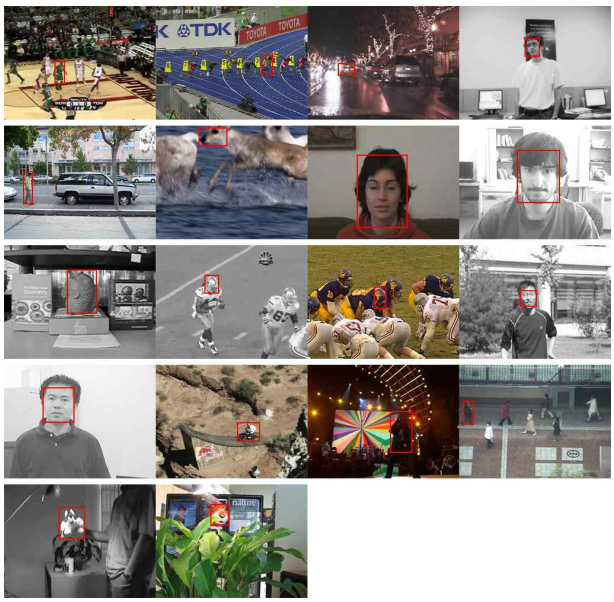

Fig. 7. Tracking sequences for evaluation

2.3 Multi-dimensional Features

MOSSE 추적기는 입력영상을 -0.5에서 0.5 사이로 밝기값을 정규화하여 사용하였다. 그리고 multi-DCF와 제안한 방법 multi-WDCF에 사용할 다차원 특징을 추출하기 위해 f-HOG[12]을 사용하였다. f-HOG 특징은 기울기 및 텍스쳐 기반 특징을 다차원으로 추출하여 준다. 실험에서는 f-HOG에서 추출한 31차원의 특징과 MOSSE 추적기에 사용되는 영상을 추가하여, 총 32차원 특징을 사용하였다.

2.4 Parameters

제안한 알고리즘에서 사용되는 변수 α 는 2로 고정하고 실험하였다. α 는 값이 커질수록 PSPR이 가장 큰 추적기의 결과와 최종결과가 유사해진다. 따라서 α가 높은 값을 가지게 되면 소수개의 추적기에 의해 표적의 위치가 결정되어 결과에 좋지 않은 영향을 준다. 반면 α가 0 이면 기존의 multi-DCF와 같은 성능을 가진다. 그렇기 때문에 실험을 통하여 최대의 성능을 나타내는 α를 결정하였다. Fig. 8은 α에 따른 DP, OP 그래프이다.

실표적과 유사도가 거의 같은 유사표적이 나타났을 때, 낮은 가중치를 가진 추적기가 추적결과에 오류를 발생시킬 확률이 높아진다. $\beta(0<\beta<1)$는 이러한 상황에서 저성능 추적기에 의한 노이즈를 제거함으로 써 오류를 줄이는데 사용한다. 따라서 일반적인 추적 상황에서는 낮은 β 값은 결과에 유의미한 영향을 주지 않는다. 하지만 α 와 마찬가지로 값이 커질수록 PSPR이 높은 소수개의 추적기에 의해 결과가 결정되기 때문에

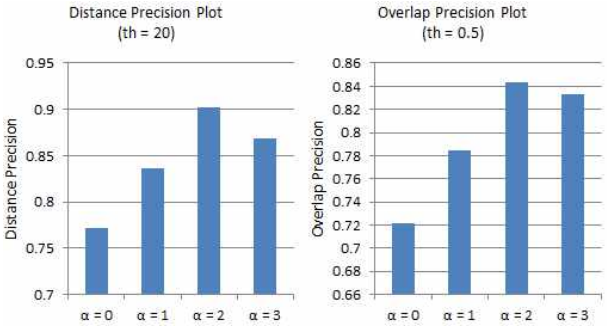

Fig. 8. Average DP/OP Precision Plot
(DP th = 20, OP th = 0.5)

높은 값을 지양해야한다. 본 논문에서는 10%의 추적기가 상관응답의 노이즈로 작용한다고 보고 β를 0.1로 설정하였다.

그리고 PSPR을 계산하기 위한 제외영역은 크기를 11×11로 설정하였다. PSR과 PSPR은 $\max_{peak}$와 sidelobe 영역의 통계수치를 이용한다는 점에서 계산 방법이 유사하기 때문에 PSR을 제시한 MOSSE에서 사용하는 제외영역의 크기를 그대로 사용하였다. 그리고 표적의 크기변화를 고려하지 않기 때문에 제외영역의 크기를 고정하였다.

3. Tracking Performance Evaluation

본 논문에서 기본 추적기로 사용하는 MOSSE 추적기와 이를 다차원으로 확장한 버전의 추적기 multi-DCF, 그리고 제안한 방법의 multi-WDCF 추적기의 성능을 비교 측정하였다.

또한 제안한 추적기 성능평가 방법인 PSPR과 PSR의 성능을 비교하기 위하여 PSR을 이용한 상관응답 조합으로 추적기를 구성하고 평가에 포함시켰다. 각 차원의 상관응답 영향도를 결정하는 식(11)의 PSPR을 PSR로 수정하였고, 나머지 설정은 동일하게 하였다. 수정한 식은 아래와 같다.

$$\pi^l = \frac{(psr^l)^\alpha}{\sum_{i=1}^{d}(psr^i)^\alpha} \tag{15}$$

3.1 CPE Performance

Table 1.는 각 실험 영상별로 각 추적기의 중심위치오차(CPE)를 보여준다. 가장 좋은 추적 성능을 보여주는 결과값을 진하게 표시하였다. 평균 중심위치오차가 약 13픽셀을 보여준 multi-WDCF-PSPR 이 가장 좋은 성능을 나타내었다.

3.2 DP Performance

주어진 실험영상을 가지고 거리 임계값을 1에서 40까지 변경하면서 거리정밀도(DP) 값을 측정하였다. 모든 실험영상의 거리정밀도를 평균한 그래프를 Fig. 9에 나타내었다. MOSSE, multi-DCF, multi-WDCF-PSR, multi-WDCF-PSPR 순으로 추적성능이 좋아지는 것을 알 수 있다.

Table 1. Per video CPE on the 18 sequences

Sequence	multi-WDCF-PSPR	multi-WDCF-PSR	multi-DCF	MOSSE
Basketball	**10.30**	175.00	74.20	72.00
Bolt	36.00	**12.00**	379.00	413.00
CarDark	2.78	2.59	**1.81**	5.21
David2	3.31	3.21	**2.39**	3.89
David3	**5.11**	88.40	98.50	92.40
Deer	4.51	4.52	**4.47**	4.85
FaceOcc1	12.80	12.60	12.70	**11.70**
FaceOcc2	5.57	5.46	**5.30**	5.63
Fish	3.47	3.32	**2.84**	40.40
Football	6.97	**6.67**	16.40	15.80
Fottball1	**5.25**	5.81	7.64	28.00
Jumping	**59.90**	84.00	85.10	88.60
Mhyang	3.01	2.76	**2.65**	3.90
M.t. Bike	6.28	**6.11**	6.53	6.26
Singer2	10.20	9.87	**8.51**	186.00
Subway	2.64	2.69	**2.43**	163.00
Sylvester	16.50	16.20	14.00	**7.92**
Tiger1	**40.40**	63.90	72.70	73.90
평균	13.06	28.06	44.29	67.91

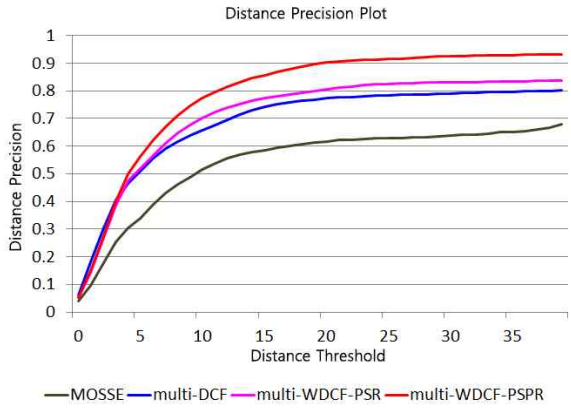

Fig. 9. Average Distance Precision Plot

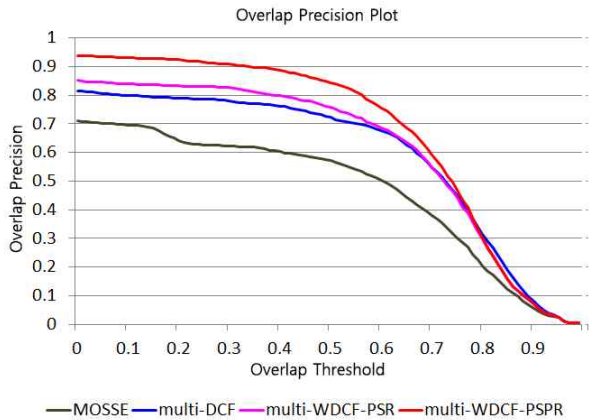

Fig. 10. Average Overlap Precision Plot

3.3 OP Performance

임계값을 0에서 1까지 변경하면서 겹침정밀도 값을 측정하였고, 모든 실험영상에 구해진 겹침정밀도를 평균하여 Fig. 10에 나타내었다. 거리정밀도와 마찬가지로 제안한 방법인 multi-WDCF-PSPR 추적기가 가장 좋은 추적성능을 보여주었다.

모든 성능지표(CPE, DP, OP)에서 PSPR로 구성한 추적기가 PSR로 구성한 추적기보다 좋은 성능을 나타내었다. 이는 PSPR 추적기 평가방법이 PSR 보다 추적의 성공여부를 더 정확하게 판단하였다는 것을 의미한다. 또한 이러한 결과를 통해서 추적기 성능평가 방법을 개선하여 추적 성능을 높일 수 있다는 것을 알 수 있다.

V. Conclusions

본 논문에서는 DCF 및 이를 확장한 multi-DCF 기반 추적방법의 한계를 극복하기 위해 새로운 상관응답 조합방법을 제시하였다. 각 특징차원의 추적기는 상관응답을 분석하여 추적성능을 평가하고, 이 값에 따라 추적결과에 기여도를 반영한다. 이를 위하여 배경의 평균 상관응답을 이용하는 PSR 추적기 신뢰도 평가 방법 대신 지역적 상관응답 분포특성을 이용하는 PSPR 추적기 신뢰도 평가 방법을 제시하였다. PSPR을 이용하여 계산된 각 차원 상관응답 가중치로 최종 상관응답을 생성하는 multi-WDCF기반 추적기를 구성하였다. 제안된 multi-WDCF-PSPR 추적기는 다양한 영상에서 하나의 특징을 이용한 MOSSE 추적기는 물론, 다차원 확장 추적기와 PSR을 이용하여 조합된 multi-WDCF-PSR 추적기보다 추적에 뛰어난 성능을 보여주었다.

한편, multi-DCF 추적기는 표적의 크기를 추정하는데 적용하여 추적성능을 높일 수 있다. 크기 공간에서 추출한 특징을 이용한다는 차이점이 있지만, 다중 상관필터를 융합하여 최종 상관필터를 구하는 과정은 표적의 이동량을 추정하는 방법과 동일하다. 따라서 제안한 추적방법 multi-WDCF는 multi-DCF를 이용한 크기 추정방법에 쉽게 적용할 수 있다. 본 논문에서는 제안 방법을 표적의 이동량을 추정하는데 만 적용하였지만, 앞으로 multi-WDCF를 표적의 크기를 예측하는데 적용하여 추적기의 성능을 높이는 연구를 진행할 예정이다.

REFERENCES

[1] A. W. M. Smeulders, R. Cucchiara, and A. Dehghan, “Visual tracking: An experimental survey,” IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 36, No. 7, pp. 1442-1468, Jul. 2014.

[2] Y. S. Jung, and H. I. Choi, "A Study on Improving the Adaptive Background Method for Outdoor CCTV Object Tracking System," Journal of The Korea Society of Computer and Information, Vol. 20, No. 7, pp. 17-24. Jul. 2015.

[3] K. H. Lee, " Face Tracking Using Face Feature and Color Information," Journal of The Korea Society of Computer and Information, Vol. 18, No. 11, pp. 167-174. Nov. 2013.

[4] M. Kristan, R. Pflugfelder, A. Leonardis, J. Matas, and et al, “The visual object tracking vot2014 challenge results,” In European Conference on Computer Vision Workshop, pp. 1-27, Zurich, Switzerland, Sep. 2014.

[5] D. S. Bolme, B. A. Draper, and J. R. Beveridge, “Average of synthetic exact filters,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2105-2112, Miami Beach, Florida, Jun. 2009.

[6] D. S. Bolme, J. R. Beveridge, B. A. Draper, and Y. M. Lui, “Visual object tracking using adaptive correlation filters,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2544-2550, San Francisco, Calif, Jun. 2010.

[7] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, “Exploiting the circulant structure of tracking-by-detection with kernels,” European Conference on Computer Vision, pp. 702-715, Springer, Berlin Heidelberg, Oct. 2012.

[8] M. Danelljan, F. S. Khan, M. Felsberg, and J. Weijer, “Adaptive color attributes for real-time visual tracking,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1090-1097, Columbus, Ohio, Jun. 2014.

[9] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, “High-speed tracking with kernelized correlation filters,” IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 37, No.3, pp. 583-596, Mar. 2015.

[10] M. Danelljan, G. Hager, F. S. Khan, and M. Felsberg, “Accurate scale estimation for robust visual tracking,” British Machine Vision Conference, pp. 65.1-65.11, Nottingham, England, Sep. 2014.

[11] Y. Wu, J. Lim, and M. H. Yang, “Online object tracking: A benchmark,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2411-2418, Portland, OR, Jun. 2013.

[12] P. Felzenszwalb, D. McAllester, and D. Ramanan, “A discriminatively trained, multiscale, deformable part

model," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8, Anchorage, AK, Jun. 2008.

Authors

Tae Eun Song received the B.S. and M.S. degrees in Electrical Engineering and Computer Science from Kyungpook National University, Korea, in 2007 and 2009, respectively

He is currently working in Hanwha Systems Co., Ltd, Yongin, Korea, from 2009. He is interested in computer vision and image processing.

Kyung Hyun Jang received the B.S., M.S. and Ph.D. degrees in Computer Science and Engineering from Hanyang University, Korea, in 2005, 2007 and 2014, respectively

In 2014, he joined the Hanwha Systems. Co., Republic of Korea, and is currently a senior researcher. He is interested in computer vision and pattern recognition.