

비의도 움직임 완화 필터 기반 동영상 안정화 Video Stabilization Based on Smoothing Filter of Undesirable Motion

김 범 수*, 임 진 주*, 홍 민 철**

Beomsu Kim*, Jinju Lim*, Min-Cheol Hong**

Abstract

Please This paper presents method of video stabilization based on detection and adaptive motion smoothing filtering of undesirable motion. The proposed algorithm consists of two stages: the detection of undesirable motion and smoothing filtering of detected undesired motion. To incorporate desired properties into the motion smoothing process, the local maximum and the local minimum are defined in a set composed of the parameters of accumulative global motion. Using the local information, the constraints on detecting undesirable motions are defined. Based on these constraints, the alpha parameter of the alpha-trimmed means filter is adjusted, so that the degree of motion smoothing in the reconstructed video sequence is controlled. The experimental results demonstrated the capability of the proposed algorithm.

요 약

본 논문은 비의도 움직임 검출 및 적응적 움직임 완화 필터를 이용한 동영상 안정화 기법을 제안한다. 제안된 기법은 비의도 움직임 검출 단계와 검출된 비의도 움직임을 완화 필터링하는 단계로 구성된다. 움직임 완화 과정에 요구되는 속성들을 활용하기 위하여 누적 전역 움직임 매개변수들로 구성된 집합의 국부 최대값과 국부 최소값을 정의한다. 국부 정보를 사용하여 비의도 움직임 검출에 필요한 제약조건을 정의하며, 제약조건들을 기반으로 하여 알파-조정 평균 필터의 알파 값을 결정하여 재구성된 동영상의 움직임 완화 정도를 제어한다. 실험 결과를 통해 제안된 방식의 성능 우수성을 입증하였다.

Key words : video stabilization, detection, alpha-trimmed means filter, undesirable motion, accumulative global motion.

* School of Electronic Engineering, Soongsil University / rhand41@ssu.ac.kr / +82-2-823-0372

★ Corresponding author

※ Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (No. 2014R1A1A2053796)

Manuscript received Mar. 27, 2015; revised June. 2, 2015 ; accepted June. 6. 2015

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

휴대용 카메라 디바이스의 저가격, 소형화 및 편리성으로 인해 관련 분야에서 휴대용 카메라 디바이스의 활용도는 다양해지고 있다. 그러나 동영상 획득 과정에서의 손 떨림 및 이동형 카메라 플랫폼 등은 획득 동영상의 진동을 유발하게 되어 시각적인 왜곡 현상을 발생하게 되고, 이로 인해 비전 기반 인식 시스템, 동영상 압축 방식 등과 같은 다양한 응용 시스템의 성능 저하를 발생하게 된다[1,2].

일반적으로 동영상 안정화 시스템은 움직임 추정, 움직임 필터링, 영상 와핑(warping)의 일련의 과정으로 구성된다[3]. 움직임 추정은 동영상의 인접된 영상

사이의 상관관계를 정의하며, 블록 매칭, 누적 투영법(integral projection), 특징(feature) 기반의 움직임 추정 기법 등의 방식이 활용되어 왔다[4,5]. 특징 기반의 움직임 추정 기법은 스케일링(scaling), 평행이동(translation) 및 회전(rotation) 등을 표현하기에 적합하며, 영상간의 전역 움직임(global motion)에 의해 움직임 궤적(trajjectory)을 효과적으로 추정할 수 있다[5,6,7]. 두 번째 단계인 움직임 완화 필터링에서는 불규칙한 움직임 현상을 제거하기 위해 손 떨림에 의한 비의도 움직임을 예측하여 움직임 완화 필터를 적용하게 된다[7,8,9,10]. 움직임 완화 필터 과정에서는 움직임 필터의 완화 강도(strength)가 미약한 경우 비의도 움직임 제거 결과가 만족스럽지 못하며, 필터 완화 강도가 커지면 의도한 움직임 성분도 제거하여 복원 동영상에 왜곡 현상을 발생시킬 수 있다. 그러므로 흔들림이 존재하는 동영상에서 비의도 움직임을 검출하고, 그 정도에 따라 움직임 완화 과정을 적용할 필요가 있다. 마지막 처리단계에서는 필터링된 매개변수들을 영상 렌더링(rendering), 호모그래피(homography) 근사화, 컨텍스트(context) 보존 와핑을 이용하여 동영상 재구성하기 위해 사용한다[7,10]. 본 논문은 비의도 움직임 제거에 초점을 두고 있으며, II절에서 비의도 움직임 제거 필터와 관련된 이전의 연구 내용들에 대해 기술한다.

본 논문에서는 비의도 움직임 검출 및 알파-보정 평균(alpha-trimmed means) 필터를 이용한 동영상 안정화 방식에 대해 제안한다. Kanade-Lucas-Tomas (KLT)[11,12]을 이용하여 추출된 특징점을 기반으로 움직임 궤적을 구성하기 위해 누적된 전역 움직임 매개 변수들로 구성된 집합을 정의하며, 어파인 변환(affine transformation) 매개변수를 예측하기 위해 random sample consensus (RANSAC)[13] 방식을 사용한다. 누적된 전역 움직임 매개변수들로 구성된 집합에서 국부 최대 및 최소값을 이용하여 비의도 움직임 매개변수들을 검출하기 위한 제약조건들을 정의하며, 이를 기반으로 적응적인 알파-보정 평균 움직임 완화 필터의 알파 매개변수를 결정하여 비의도 움직임의 완화 정도를 제어한다.

본 논문은 다음과 같이 구성된다. II절에서는 전역 움직임 추정 기법과 움직임 완화 필터에 대한 기존의 연구들에 대해 간략히 설명한다. III절에서는 제안된 알파-보정 움직임 완화 필터에 대해 기술한다. 움직임 궤적을 나타내는 집합으로부터 비의도 움직임 성분을 검출하기 위한 제약 조건을 설정하고, 상기 제약 조건을 이용하여 비의도 성분 제거를 위한 알파-보정 필터의 알파 매개변수 결정 과정에 대해 설명한다. IV장 및 V절은 실험 결과 및 결론에 대하여 기술

한다.

II. 관련 연구

전역 움직임 추정 방식과 관련된 방식들 중, 블록 매칭 움직임 추정과 유사 움직임 모델 방식을 결합시킨 방식은 연속된 두 영상 사이에 선형 이동만이 존재한다는 가정 하에 제안되었다[4]. 또한, optical flow 과 최소 제곱 (least squares)을 이용한 어파인(affine) 움직임 모델을 결합하여 전역 움직임을 유도하는 방식도 제안되었다[14]. 상기 방식들은 움직임의 평행 이동, 회전, 확대 및 축소와 같은 다양한 움직임 유형을 처리하기에는 성능의 한계가 있다. 이와 같은 문제점을 해결하기 위해 특징 기반의 전역 움직임 추정 기법과 관련된 연구가 활발히 진행되었다. 예를 들어, 스케일 불변 특징 변환 (SIFT: Scale Invariant Feature Transform)에서는 descriptor를 사용하여 특징점간의 연계를 정의하여, 스케일, 회전 및 변화에 강인한 특성을 지닌 방식이 제안하였다[8]. 또한 SIFT기반의 전역 움직임 추정 방식의 계산량을 절감시키기 위한 Speed Up Robust Features (SURF) 방식이 발표되었다[6]. 최근에는 특징점을 기반으로 하는 방식으로는 파티클(particle) 필터와 SIFT를 결합한 하이브리드 방식[5], 호모그래피 특징 추적 모델[10], KLT 기법과 결합된 3-D모델[15] 등이 활발히 연구되었다.

동영상 안정화 시스템에서 움직임 완화 필터의 목적은 비의도 움직임 성분을 제거하는 것으로서 이와 관련된 다양한 필터 방식들이 연구되었다. 초기 연구로서 연속된 영상간의 예측한 optical flow를 보정하기 위해 어파인 움직임 모델 기반의 정규화 움직임 완화 필터가 소개되었다[1]. CMOS (Complementary Metal Oxide Semiconductor)에서 어파인 전역 움직임 왜곡 현상을 제거하기 위한 저 연산량의 선형 움직임 완화 필터 방식이 제안되었다[2]. 그러나 상기 방식들은 필터링 과정에서 선형 움직임 모델을 가정하여, 실제 움직임이 비선형적일 경우 복원 동영상에 심각한 왜곡 현상을 유발하는 문제점이 있다. 이와 같은 문제를 해결하기 위해 다양한 적응적 움직임 완화 필터들이 최근에 연구되었다. 특히, SIFT 방식을 사용하여 추정된 전역 움직임 매개변수들의 불규칙적인 산란현상을 제거하기 위해서 파티클 필터 방식[8], 칼만 필터 기반의 완화 필터 [7], 고차 B-스플라인 움직임 완화 필터[9], 움직임 궤적 기반의 저역 통과 필터[10] 등의 다양한 적응 방식들이 연구되어 왔다.

III. 비의도 움직임 완화 필터 기반 동영상 안정화

연속된 영상간의 관계는 추출된 특징점을 기반으로 하는 전역 움직임에 의해 표현될 수 있으며, 특정 추출의 정확도는 영상 안정화의 문제를 해결하는데 중요한 역할을 한다. KLT방식은 영상의 특징점을 효과적으로 추출할 수 있음이 입증되었으며, 특징들 사이의 관계는 다음과 같이 기술될 수 있다.

$$\mathbf{p}_t = \mathbf{H}_t \mathbf{q}_{t-1} = \mathbf{H}_t [\mathbf{p}_{t-1}^T, 1]^T. \quad (1)$$

위 식에서 $\mathbf{p}_t = [x_t, y_t]^T$ 는 임의의 영상 f_t 에서 특징의 위치를 나타내고 $\mathbf{p}_{t-1} = [x_{t-1}, y_{t-1}]^T$ 은 이전 영상 f_{t-1} 에서 대응되는 특징의 위치를 나타낸다. 식 (1)에서 T 는 벡터의 치환 (transpose) 연산을 의미하며, \mathbf{p}_t 와 \mathbf{p}_{t-1} 사이의 관계는 어파인 변환 행렬 \mathbf{H}_t 를 이용하여 다음과 같이 표현될 수 있다.

$$\begin{aligned} \mathbf{H}_t &= \begin{bmatrix} \phi_t(1,1) & \phi_t(1,2) & \phi_t(1,3) \\ \phi_t(2,1) & \phi_t(2,2) & \phi_t(2,3) \end{bmatrix} \\ &= \begin{bmatrix} S_t \cos \theta_t - S_t \sin \theta_t & T_{t,x} \\ S_t \sin \theta_t & S_t \cos \theta_t & T_{t,y} \end{bmatrix}. \end{aligned} \quad (2)$$

식 (2)에서 S_t 와 θ_t 는 각각 \mathbf{p}_t 와 \mathbf{p}_{t-1} 사이의 스케일 및 회전 변수를 의미하며, $T_{t,x}$ 와 $T_{t,y}$ 는 특징의 수평, 수직 방향으로의 평행이동을 나타낸다. 본 연구에서는 식 (2)의 각 매개변수들은 예측하기 위해 RANSAC 기법을 이용하였다.

본 논문은 예측된 매개변수들을 이용하여 비의도 움직임 성분을 검출하고, 검출된 비의도 성분을 제거하는 방식을 제안한다. 움직임 궤적을 표현하기 위해 식 (2)의 매개변수들로 구성된 집합을 정의하고, 정의된 집합을 이용하여 비의도 성분을 검출하기 위한 제약 조건을 설정한다. 제약 조건에 의해 검출된 비의도 움직임 성분의 통계 특성을 활용하여 알파-보정 움직임 완화 필터의 알파 매개변수를 조정함으로써 재구성 동영상의 움직임 완화 정도를 제어하는 방식에 대해 제안한다.

1. 비의도 움직임 검출

\mathbf{H}_t 의 (i,j) 번째 매개변수가 비의도 움직임 성분을 포함하는 경우 왜곡 모델은 다음과 같이 기술될 수

있다.

$$\phi_t(i,j) = \phi_{t,desired}(i,j) + \phi_{t,undesired}(i,j). \quad (3)$$

위 식에서 $\phi_{t,desired}(i,j)$ 및 $\phi_{t,undesired}(i,j)$ 는 각각 (i,j) 번째의 어파인 변환 매개변수의 의도 및 비의도 움직임 성분을 나타낸다. 일반적으로, 비의도 움직임에 대한 정보는 제한적이기 때문에 직접 추출하여 제거하는 것은 매우 어렵다.

일반적으로 의도된 움직임은 인접 영상간의 특징들이 완전한 형태로 표현되며, 비의도 움직임은 불규칙한 형태의 특성을 갖고 있다[10]. 상기와 같은 의도 및 비의도 움직임에 대한 특성을 활용하기 위해 본 논문에서는 움직임 궤적 정보를 이용하여 비의도 움직임 성분을 검출하기 위한 제약 조건을 설정한다. 움직임 궤적은 다양한 형태로 표현될 수 있으며, 본 연구에서는 누적 어파인 변환 매개변수 (AATP: accumulative affine transformation parameter)들을 움직임 궤적을 표현하기 위해 사용한다. 각각의 어파인 변환 매개변수가 다른 매개변수들과 상관성이 없다고 가정하여 첫 번째 프레임부터 현재 프레임까지 (i,j) 번째 누적 평면 변환 매개변수는 다음과 같이 표현될 수 있다.

$$\Phi_t(i,j) = \sum_{m=1}^t \phi_m(i,j). \quad (4)$$

식 (4)를 이용하여 (i,j) 번째 어파인 변환 매개변수의 궤적을 나타내는 집합을 다음과 같이 정의한다.

$$S_t(i,j) = \{k \in \{t-L, \dots, t\} | \Phi_k(i,j)\}. \quad (5)$$

위 식에서 $(L+1)$ 은 누적 어파인 변환 매개변수 집합(AATPS: AATP set)의 크기를 나타낸다. 이미 언급한 바와 같이, 집합 내에서 일정한 변화도를 초과하는 원소의 수가 감소할수록 비의도 움직임의 정도가 감소하며, 비의도 움직임의 정도가 심해질수록 원소들의 변화량은 커지게 된다. 이와 같은 특성을 반영하기 위해 집합의 원소들의 변화에 따라 다음과 분류 과정을 수행한다.

$$ds_k = \begin{cases} 0 & \text{if } T_1 < \Phi_k(i,j) < T_2 \\ 1 & \text{otherwise} \end{cases}. \quad (6)$$

식 (6)에서 $\Phi_k(i,j) \in S_t(i,j)$ 이며, $ds_k = 1$ 은 $\Phi_k(i,j)$ 가 비의도 움직임 성분을 포함하는 것을 의미한다. 반면에, $ds_k = 0$ 은 $\Phi_k(i,j)$ 가 다른 원소들과 비슷하거나 비의도 움직임 정도가 매우 작은 경우를 나타낸다. 더불어, T_1 와 T_2 는 임계값을 의미하며, 다음과 같이 설정된다.

$$\begin{aligned} T_1 &= (\Phi_{\max} - \Phi_{\min}) \times c_1 + \Phi_{\min}, \\ T_2 &= (\Phi_{\max} - \Phi_{\min}) \times c_2 + \Phi_{\min}, \end{aligned} \quad (7)$$

위 식에서 c_1 과 c_2 ($c_1 < c_2$)는 상수를 나타내고, Φ_{\max} 와 Φ_{\min} 은 다음과 같이 정의된다.

$$\begin{aligned} \Phi_{\max} &= \max_{\Phi_l(i,j) \in S(i,j)} \Phi_l(i,j), \\ \Phi_{\min} &= \min_{\Phi_l(i,j) \in S(i,j)} \Phi_l(i,j). \end{aligned} \quad (8)$$

고정된 임계값을 사용하는 경우, 식 (6)의 검출 과정의 성능은 임계값에 매우 민감한 특성을 갖기 때문에 누적 매개변수에 대한 국부 정보와 같은 유용한 특성이 검출 과정에 결합되어야 한다. 그러므로 본 논문에서는 집합의 국부 최대 및 최소값을 활용하여 식(6)의 비의도 검출기의 제약 조건으로 사용한다. 더불어, 집합 내의 원소들의 변화 정도가 증가함에 따라 비의도 성분이 커지므로 AATPS의 변화 정도를 다음과 같이 정의하였다.

$$M_t = \sum_{k=t-L+1}^t ds_k. \quad (9)$$

의도된 움직임이 완만하게 변화되는 특성을 갖고 있는 관계로 M_t 가 클수록 집합 내에 비의도 움직임이 많은 것으로 추정할 수 있다. 위의 비의도 검출 결과를 이용하여 변화 정도에 따라 움직임 완화 필터의 강도를 제어하여 비의도 움직임 성분을 제거하고자 한다.

1. 적용적인 알파 보정 평균 필터

비의도 움직임 성분은 일반적으로 비선형적인 특성을 갖기 때문에 일반적인 선형 움직임 완화 필터를 활용하여 성능 향상을 추구하기에는 한계성이 있다. 비선형 필터는 영상의 노이즈 제거 문제를 해결하기 위해 오랜 기간 연구되어 왔으며, 알파-보정 평균 필터는 알파 매개변수를 제어하여 저연산량으로 비선형

특성을 반영할 수 있는 장점을 지니고 있다[16].

본 논문에서 알파-보정 평균 필터를 AATPS에 적용하기 위해서 $S_t(i,j)$ 를 다음과 같이 오름차순으로 정렬한다.

$$\tilde{S}_t(i,j) = \{k \in \{t-L, \dots, t\} | \tilde{\Phi}_k(i,j)\}. \quad (10)$$

위 식에서 $\tilde{\Phi}_{t-L}(i,j) \leq \tilde{\Phi}_{t-L+1}(i,j) \leq \dots \leq \tilde{\Phi}_t(i,j)$ 이며, 알파-보정 평균 필터 결과는 (i,j) 번째 완화된 누적 매개변수를 이용하여 다음과 같이 얻을 수 있다.

$$\hat{\Phi}_t(i,j) = \frac{1}{L+1-2\alpha} \sum_{k=t-L+\alpha}^{t-\alpha} \tilde{\Phi}_k(i,j). \quad (11)$$

식 (11)에서 필터링 결과의 완화 정도는 매개변수인 α 에 의해 결정 된다[17]. 예를 들어, α 가 '0'일 때 평균 (mean) 필터의 역할을 하며, $\alpha = L/2$ 일 때는 미디어 (median) 필터의 역할을 하게 된다. 움직임 완화 문제에서 비의도 움직임은 노이즈 제거의 첨부 노이즈로 간주될 수 있으므로 비의도 움직임 성분이 증가할수록 움직임 완화 필터의 강도가 커져야 한다. 그러므로 α 는 다음과 같은 특성을 가져야 한다.

$$\alpha \propto \frac{1}{M_t}. \quad (12)$$

식 (12)의 관계를 만족하는 알파 매개변수는 다양한 방법으로 정의될 수 있다. 그러나 비선형 특성을 갖는 변화량과 알파 매개변수의 상관관계를 정확하게 만족시키는 함수를 유도하는 것은 매우 어렵다. 본 논문에서는 다음과 같이 알파 매개변수를 결정하였다.

$$\alpha = \begin{cases} \alpha_1 & \text{if } M_t > Thr \\ \alpha_2 & \text{otherwise} \end{cases}. \quad (13)$$

위 식에서 $\alpha_1 < \alpha_2$ 이고, Thr 는 임계값을 의미한다.

식 (11)은 비의도 움직임을 제거하는 누적 어파인 변환 매개변수를 의미하므로 (i,j) 번째 필터링 어파인 변환 매개변수는 다음과 같이 구할 수 있다.

$$\begin{aligned} \bar{\phi}_t(i, j) &= \phi_t(i, j) + \hat{\phi}_t(i, j) - \phi_t(i, j) \quad (14) \\ &\approx \phi_t(i, j) + \left(\sum_{k=1}^{t-1} \phi_k(i, j) + \phi_{t,desired}(i, j) \right) \\ &\quad - \left(\sum_{k=1}^{t-1} \phi_k(i, j) + \phi_t(i, j) \right) \\ &= \hat{\phi}_{t,desired}(i, j) \end{aligned}$$

식 (14)를 통해 얻어지는 매개변수들을 이용하여 비의도 움직임 성분이 완화된 어파인 변환 행렬을 다음과 같이 표현할 수 있다.

$$U_t = \begin{bmatrix} \bar{\phi}_t(1,1) & \bar{\phi}_t(1,2) & \bar{\phi}_t(1,3) \\ \bar{\phi}_t(2,1) & \bar{\phi}_t(2,2) & \bar{\phi}_t(2,3) \end{bmatrix} \quad (15)$$

위 식을 활용하여 비의도 움직임 성분을 제거한 영상 \bar{f}_t 는 현재 영상 f_t 의 완화된 움직임 매개변수에 따라 양방향 보간(bilinear interpolation) 및 와핑 과정을 이용하여 다음과 같이 재구성하게 된다[18].

$$\bar{f}_t \leftarrow f_t(U_t) \quad (16)$$

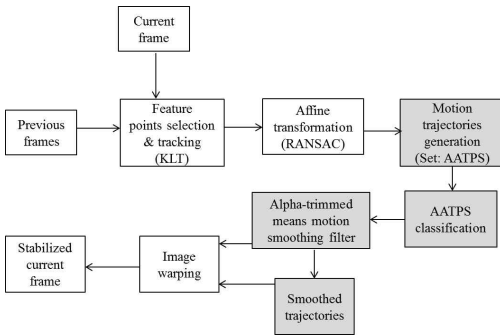


Fig. 1. Flowchart of proposed algorithm
그림 1. 제안된 방식의 흐름도

그림 1에 제안 방식을 회색박스로 표현하여 나타난 전체 시스템의 흐름도를 도시하였다. 본 논문에서는 의도 및 비의도 움직임에 대한 사전 정보 없이 적응적으로 움직임의 완화 정도를 제어함으로써 비의도 움직임을 효과적으로 제거하는 방식에 대해 제안하였다.

IV. 실험결과

제안된 방식을 평행 이동, 회전, 스케일링 및 조도 변화 등 다양한 환경을 포함하는 그림 2와 같은 “Hippo”, “Outdoor”, 및 “Zoom” 동영상을 사용하여 실험하였다. 실험에 사용한 동영상은 초당 30 프레임 (fps: frames per second)이다. 각 동영상의 특성은 표 1에 간략하게 기술하였다.

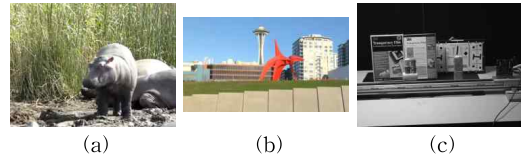


Fig. 2 Test video sequences: (a) “Hippo”, (b) “Outdoor”, (c) “Zoom”

그림 2. 실험에 사용한 동영상: (a) “Hippo”, (b) “Outdoor”, (c) “Zoom”

Table 1. Test video characteristics

표 1. 테스트 동영상 특성

Video seq.	Spatial resolution	Total number of frames	Video properties
Hippo	640×480	429	translation with stationary camera
Outdoor	640×360	496	translation and rotation with non-stationary camera
Zoom	672×512	270	Zoom-in/out and illuminance change with stationary camera

본 논문에서 제안된 움직임 제적 기반의 알파-보정 평균 필터 (ATMMT) 방식은 움직임 차이를 이용한 테일러급수 (TSMD) 방식[3], 움직임 제적을 이용한 칼만 필터 (KFMT) 방식[7] 및 움직임 제적을 이용한 적응 필터 (AFMT) 방식[10]과 성능 비교를 하였다. 성능 평가를 위해 영상간의 변환 충실도 (ITF: Inter-frame Transformation Fidelity)를 사용하였다. ITF는 연속된 두 영상간 변화의 완만도를 나타내며, 영상간 변화가 작을수록 ITF는 커지게 된다[19]. 총 영상수가 R 인 동영상의 ITF는 다음과 같이 정의된다.

$$ITF = \frac{1}{R-1} \sum_{r=2}^R PSNR(r), \quad (17)$$

위 식에서 최대 신호 대 잡음 비 (PSNR: Peak-to-signal ratio)는 다음과 같이 정의된다.

$$PSNR(r) = 10 \log_{10} \frac{f_{\max}^2}{\frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N [f_r(x,y) - f_{r-1}(x,y)]^2} \quad (18)$$

위 식에서 M 및 N 은 수평 및 수직 방향의 공간 해상도를 나타내고, f_{\max} 는 최대 밝기 값을 나타낸다.

식 (5)에서 AATPS의 원소의 수가 많을수록 $\Phi_{t-L}(i,j)$ 와 $\Phi_t(i,j)$ 의 상관관계가 낮아지게 되어 $\Phi_{t-L}(i,j)$ 의 정보는 $\Phi_t(i,j)$ 에 대한 국부 정보로 사용하기에 적합하지 않게 된다. 반면, AATPS의 원소의 개수가 작아질수록 비의도 움직임 검출과 필터링 과정에 사용되는 제약 조건 설정의 신뢰성이 떨어지게 된다. 본 실험에서, $L=60$ 일 때, 국부 통계 특성이 충분히 반영되는 것이 관찰되었다.

식 (7)의 비의도 움직임 검출 과정에서 사용되는 상수 c_1 이 증가함에 따라 더 엄격한 범위가 적용되어 상대적으로 많은 AATPS의 원소들이 비의도 움직임 성분을 포함하는 것으로 판별된다. 이와 같은 경우 움직임 완화 필터의 강도가 커지게 되어 의도적인 움직임 정보도 완화되어 재구성된 영상에 왜곡 현상을 유발하게 된다. 더불어, c_2 가 증가할수록 비의도 검출 범위가 커지게 되어 비의도 움직임으로 검출되는 원소의 수가 감소하게 된다. 상기와 같은 경우, 비의도 움직임 완화 필터의 강도가 작아지게 되어 비의도 움직임 제거가 효과적으로 못하다. 실험을 통해 $0.05 \leq c_1 \leq 0.1$ 와 $0.9 \leq c_2 \leq 0.95$ 에서 가장 효과적인 성능을 나타냄을 확인할 수 있었으며, 본 실험에서는 $c_1 = 0.05$, $c_2 = 0.95$ 로 설정하였다.

재구성된 동영상의 움직임 완화 정도는 알파 매개변수 및 식(13)에서 설정된 임계치에 의해 결정 된다. 알파 매개변수가 작은 경우에는 움직임 완화 필터의 필터 강도가 크게 설정되고, 임계치가 낮아짐에 따라 동일한 비의도 움직임 원소 수에 대해 상대적으로 필터 강도가 크게 설정된다. 실험의 결과에 기초하여 $\alpha_1 = 0, \alpha_2 = 7, Thr = 40$ 일 때 가장 좋은 결과를 얻을 수 있음을 확인할 수 있었다.

표 2에 다양한 동영상에 대한 ITF 성능을 비교하였다. 본 논문에서 제안된 ATMMT 방식이 이동, 회

전등의 다양한 움직임을 갖는 동영상에 대해서 지속적으로 다른 방식보다 우수한 성능을 갖고 있음을 확인할 수 있다. 특히 전역 움직임이 정적인 환경에서 미의도 움직임이 존재하는 영상에 대한 ITF 이득은 상대적으로 높은 것으로 확인하였다. 그러나 “Zoom” 동영상인 경우 TSMD 방식이 다른 방식보다 성능 향상이 우월함을 알 수 있다. 이와 같은 이유로는 TSMD 방식이 정적 환경에서 선형 움직임 모델을 기반으로 비선형 움직임을 근사화하는 방식을 사용하여, zoom-in/out과 같은 비선형 움직임을 복원 영상에 반영하기에는 한계가 존재한다. 이와 같은 이유로 재구성 동영상에서 zoom-in/out 전환 영상들이 유사하게 표현되어 상대적으로 높은 ITF 값을 얻게 된다. 이와 같은 복원 영상의 왜곡 현상은 시각적 성능 비교에서 확인할 수 있다.

Table 2. ITF comparisons (unit: dB)

표 2. ITF 비교 (단위: dB)

	Observed	TS MD	KF MT	AF MT	ATM MT
Hippo	19.65	21.06	21.59	20.61	23.51
Out door	17.95	20.01	20.56	19.75	21.05
Zoom	25.07	32.70	30.26	27.79	30.52

Table 3. APTPF comparisons (unit: msec.)

표 3. APTPF 비교 (단위: msec.)

	TS MD	KF MT	AF MT	ATM MT
Hippo	107.81	97.86	66.97	78.44
Outdoor	80.49	80.62	47.97	61.67
Zoom	136.91	112.11	65.36	75.95

각 방식의 계산량 성능 비교를 하기 위해 3.4GHz의 CPU에서 프레임당 평균 연산 시간(APTPF: Average Processing Time Per Frame)을 사용하였다. 표 3에 나타낸 바와 같이 AFMT 방식은 단순한 움직임 완화 필터를 사용한 관계로 모든 영상에 다른 방식들보다 계산량면에서 우월한 성능을 보이고 있다. 반면에 TSMD 및 KFMT 방식은 필터 처리과정에서 움직임 차이를 최소화할 때의 복잡성으로 인해 상대적으로 많은 연산량을 필요로 한다. 본 논문에서 제안된 ATMMT 방식은 AFMT 방식보다 평균 15% 높은 연산량을 필요로 한다. 그러나, ATMMT 방식은 ITF 측면에서 의도된 움직임 형태에 관계없이 AFMT 방식보다 뛰어난 성능을 보이고 있다. 더불어, ATMMT 방식은 ITF와 APTPF 모든 평가 측면에서

TSMD 및 KFMT 방식보다 우수한 성능을 갖고 있음을 확인할 수 있었다.

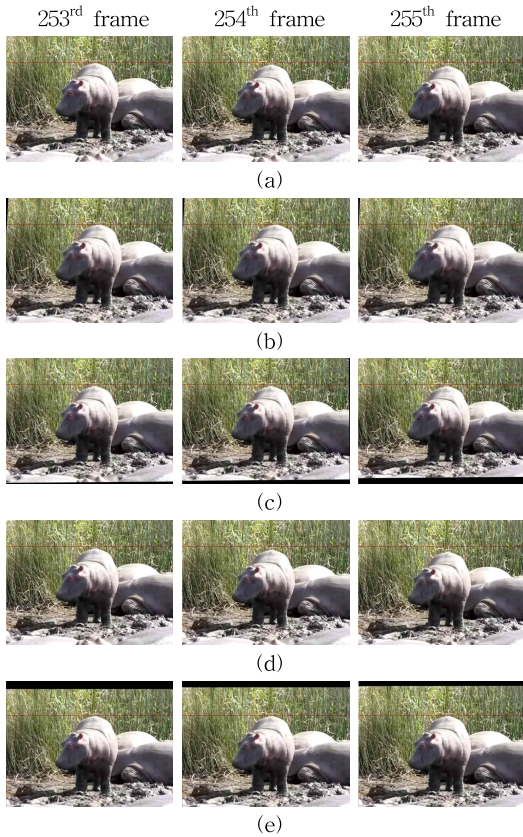


Fig. 3. Visual performance comparisons with “Hippo” sequence : (a) observed sequence, (b) result with TSMD, (c) result with KFMT, (d) result with AFMT, (e) result with ATMMT

그림 3. “Hippo”시퀀스를 이용한 시각적 성능 비교 : (a) 관찰 동영상, (b) TSMD 결과, (c) KFMT 결과, (d) AFMT 결과, (e) ATMMT 결과

그림 3의 (a)-(e)에 비의도 흔들림 현상이 존재하는 “Hippo” 동영상, TSMD 방식을 이용하여 재구성된 영상, KFMT 방식을 이용하여 재구성된 영상, AFMT 방식을 이용하여 재구성된 영상 및 ATMMT 방식으로 재구성된 영상을 나타내었다. 위의 결과에서 검은색 영역은 와핑 과정에서 정의되지 않은 누락 영역을 나타낸다. 비의도 흔들림 현상의 제거 성능을 비교하기 위해 붉은색으로 표현된 가로선을 재구성

영상들에 첨부하였다. 실험 결과를 통해 TSMD 및 AFMT 방식에서는 비의도 움직임 현상이 존재하는 것을 알 수 있다. 반면에, KFMT 및 ATMMT 방식은 비의도 움직임 현상을 효과적으로 제거하였음을 확인할 수 있었다. 그림 4의 (a)-(e)는 “Outdoor” 동영상의 비의도 움직임 존재하는 관찰 영상, TSMD, AFKF, AFMT 및 ATMMT 방식을 이용하여 재구성된 영상들을 표시하였다 (동영상 성능 비교: <http://www.youtube.com/watch?v=qVnHsVUmiPE>).

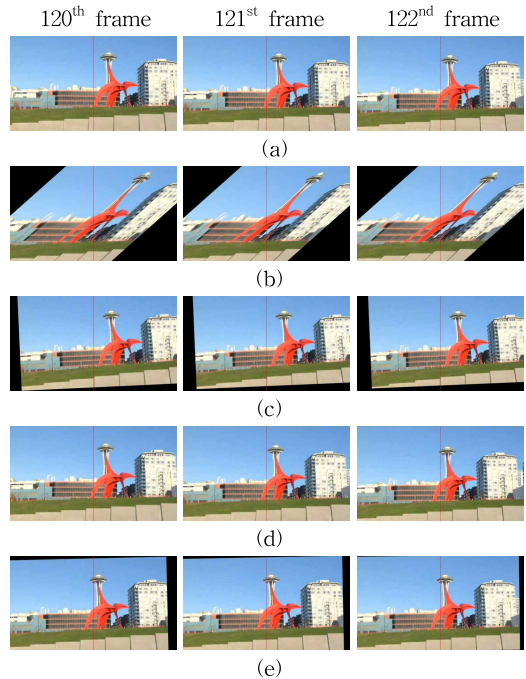


Fig. 4. Visual performance comparisons with “Outdoor” sequence : (a) observed sequence, (b) result with TSMD, (c) result with KFMT, (d) result with AFMT, (e) result with ATMMT.

그림 4. “Outdoor”시퀀스를 이용한 시각적 성능 비교 : (a) 관찰 동영상, (b) TSMD 결과, (c) KFMT 결과, (d) AFMT 결과, (e) ATMMT 결과

실험 결과를 통해 TSMD 방식은 “절단” (shearing) 왜곡 현상이 존재하는 것을 알 수 있다. 이는 선형 근사 모델을 사용되어 특징점의 회전을 재구성 영상에서 유지하지 못했기 때문인 것으로 분석된다. 더불어, KFMT 및 AFMT 방식을 이용한 결과는 상대적으로 ‘절단’ 왜곡 현상을 감소시켰음을 알 수 있으나, 왜곡 현상이 일부 존재함을 확인하였다.

반면에 ATMMT 방식은 왜곡 현상 없이 효과적으로 동영상의 복원하였음을 알 수 있다. 그림 5의 (a)-(e)를 통해 AFMT, KFMT 및 ATMMT 방식들은 “Zoom” 동영상의 zoom-in/out 현상을 유지하면서 효과적으로 흔들림 현상을 제거하였음을 알 수 있다. 그러나, 선형 움직임 근사화 모델을 사용한 TDMS 방식은 복원 영상에서 zoom-in/out 현상 보존을 못하여 연속된 영상이 유사하게 표현됨을 확인할 수 있었다.

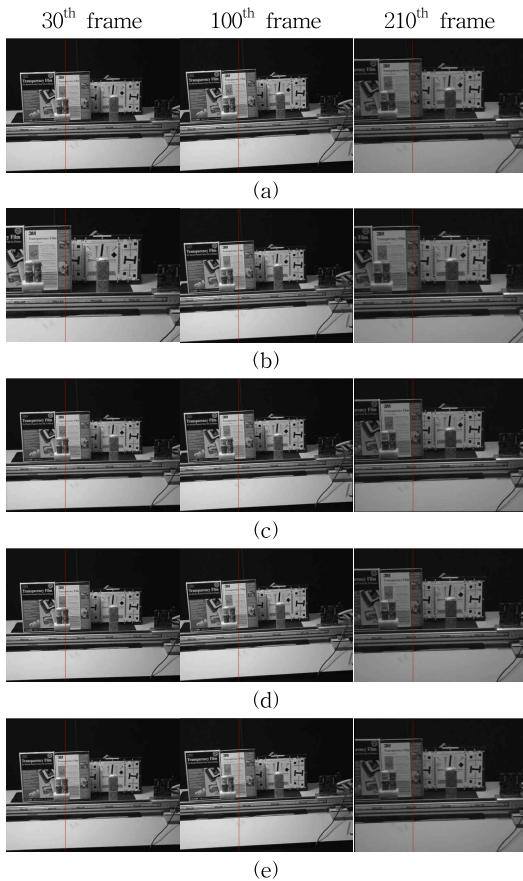


Fig. 5. Visual performance comparisons with “Zoom” sequence : (a) observed sequence, (b) result with TSMD, (c) result with KFMT, (d) result with AFMT, (e) result with ATMMT.

그림 5. “Zoom”시퀀스를 이용한 시각적 성능 비교 : (a) 관찰 동영상, (b) TSMD 결과, (c) KFMT 결과, (d) AFMT 결과, (e) ATMMT 결과

그림 6의 (a) 및 (b)에 제안된 방식을 “Hippo” 동영상

상에 적용하기 전, 후의 수평 및 수직 방향의 누적 전역 움직임 결과를 나타내었다. 위의 결과를 통해 제안된 방식을 사용하여 바람직하지 않은 비의도 움직임을 제거하였으며, 영상간의 불연속 움직임을 효과적으로 감소시켰음을 알 수 있다.

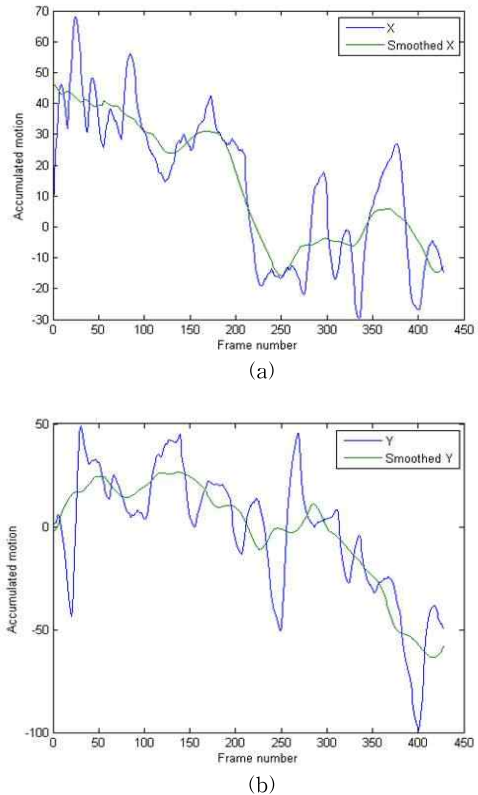


Fig. 6. Comparisons before and after motion smoothing in “Hippo” sequence : (a) in x-axis, (b) in y-axis.

그림 6. “Hippo” 동영상의 움직임 필터 적용 전,후 결과 비교 : (a) X축 방향, (b) Y축 방향.

상기 실험 결과들을 통해 제안된 방식이 밝기 변화나 의도된 움직임 형태에 관계없이 낮은 계산 비용으로 완화된 움직임 영상을 재구성 할 수 있었음을 확인할 수 있었다. 본 논문은 비의도 움직임 및 의도 움직임에 대한 사전 정보 없이, 누적 움직임 매개변수들의 국부 정보를 이용하여 제약 조건을 설정함으로써 비의도 움직임 검출 및 알파-보정 평균 필터의 알파 매개변수를 선정하는 방식에 대해 제안하였다. 실험 결과를 통해 제안된 방식의 우수성을 확인할 수 있었다.

IV. 결론

본 논문에서는 비의도 움직임 검출 및 적응적인 알파-보정 평균 필터를 기반으로 하는 동영상 안정화 방식에 대해 제안하였다. 움직임 궤적을 표현하기 위해 누적 어파인 움직임 매개변수들로 구성된 집합을 정의하였다. 정의된 집합에 대한 정보를 사용하여 비의도 움직임 검출을 위한 제약 조건을 정의하였으며, 제약 조건을 이용하여 알파-보정 평균 필터의 알파 매개변수를 선정하는 방식에 대해 기술하였다. 비의도 움직임 정도에 따라 알파 매개변수를 선정하여 재구성된 영상의 원화 정도를 제어하였다. 실험 결과를 통해 제안된 방식이 영상의 종류에 관계없이 비의도 움직임 성분을 효과적으로 감소시킬 수 있었음을 확인할 수 있었다.

현재 알파 매개변수의 선택과 관련된 추가적인 연구가 진행 중에 있으며, 보다 정교한 상관관계를 유도하여 개선된 성능을 얻을 수 있을 것으로 기대된다.

References

- [1] H. C. Chang, S. H. Lai, and K. R. Lu, "A robust real-time video stabilization algorithm," *J. Vis. Comm. Image Represent.*, vol. 17, no. 3, pp. 659 - 673, Jun. 2006.
- [2] W. H. Cho and K. S. Hong, "Affine motion based CMOS distortion analysis and CMOS digital image stabilization," *IEEE Trans. Consumer Electron.*, vol. 53, no. 3, pp. 833 - 841, Aug. 2007.
- [3] P. Rawat and J. Singhai, "Review of motion estimation and video stabilization techniques for hand held mobile video," *Signal and Image Processing: An International Journal*, vol. 2, no. 2: 159 - 168, Jun. 2011.
- [4] G. Puglisi, S. Battiato, "A robust image alignment algorithm for video stabilization purposes," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 10, pp. 1390-1400, Oct. 2011.
- [5] C. Song, H. Zhao, W. Jing, and H. Zhu, "Robust video stabilization based on particle filtering with weighted feature points," *IEEE Trans. Consumer Electron.*, vol. 58, no. 2, pp. 570-577, May 2012.
- [6] K.-Y. Huang, Y.-M. Tsai, C.-C. Tsai, and L.-G. Chen, "Video stabilization for vehicular applications using SURF-like descriptor and KD-tree," *IEEE Int. Conf. Image Processing*, pp. 3517 - 3520, Sep. 2010.
- [7] Y. G. Ryu and M. J. Chung, "Robust online digital image stabilization based on point-feature trajectory without accumulative global motion estimation," *IEEE Signal Processing Letters*, vol. 19, no. 4, pp. 233 - 265, Apr. 2012.
- [8] J. Yang, D. Schonfeld, and M. Mohamed, "Robust video stabilization based on particle filter tracking of projected camera motion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 7, pp. 945-954, Jul. 2009.
- [9] Y. Wang, R. Chang, T. W. Chua, K. Leman, N. T. Pham, "Video stabilization based on high degree B-spline smoothing," *Int. Conf. Pattern Recognition (ICPR)*, pp. 3152- 3155, Dec. 2012.
- [10] J. Dong, Y. Xia, Q. Yu, A. Su, and W. Hou "Instantaneous video stabilization for unmanned aerial vehicles," *J. Electron. Imaging*, vol. 23, no. 1, 013002, Jan. 2014.
- [11] J. Shi and C. Tomasi, "Good features to track," *Int. Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 593-600, June 1994.
- [12] C. Tomasi and T. Kanade, "Detection and tracking of point features," Carnegie Mellon University Technical Report, CMU-CS-91-132, Apr. 1991.
- [13] M. A. Fischler and R. C. Bolles, "Random Sample Consensus: A paradigm for model fitting with application to image analysis and automated cartography," *Comm. of the ACM*, vol. 24, no. 6, pp. 381-395, June 1981.
- [14] N. Ejaz, W. Kim, S. I. Kwon, S. W. Baik, "Video stabilization by detecting intentional and unintentional camera motions", 2012 Third Int. Conf. Intelligent Systems, Modelling and Simulation (ISMS), pp. 312-316, Feb. 2012.
- [15] Y.-S. Wang, F. Liu, P.-S. Hsu, and T.-Y. Lee, "Spatially and temporally optimized video stabilization", *IEEE Trans. Visual. and Comp. Graph.*, vol. 19, no. 8, pp. 1353-1361, Aug. 2013.
- [16] P. Milanfar, "A tour of modern image filtering," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 106-128, Jan. 2013.
- [17] J. Bendat and T. Watt, "Alpha-trimmed means and their relationship to median filters," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 32, no. 1, pp. 145-153, Feb. 1984.

- [18] M. Okade and P. Biswas, "Fast video stabilization in the compressed domain," IEEE Int. Conf. Multimedia and Expo, pp. 1015-1020, July 2012.
- [19] C. Morimoto and R. Chellappa, "Evaluation of image stabilization algorithms," IEEE Int. Conf. Acoustics, Speech and Signal Processing, pp. 2789-2792, May 1998.

BIOGRAPHY

Beomsu Kim (Member)



2007 : BS degree in School of Electronic Engineering, Soongsil University.
 2009 : MS degree in School of Electronic Engineering, Soongsil University.
 2010~2012 : Associate Researcher,

Korea Internet & Security Agency

2012~ : PhD candidate, School of Electronic Engineering, Soongsil University.

Jinju Lim (Student Member)



2015 : BS degree in School of Electronic Engineering, Soongsil University.
 2015~ : MS degree in School of Electronic Engineering, Soongsil University.

Min-Cheol Hong (Member)



1988 : BS degree in Electronic Engineering, Yonsei University.
 1990 : MS degree in Electronic Engineering, Yonsei University.
 1997 : PhD degree in Electrical Engineering, Northwestern University.

1990~1992 : Research Engineer, LG Electronics.

1997~1998 : Research Fellow, Northwestern University.

1998~2000 : Senior Research Engineer, LG Electronics.

2000~ : Professor, School of Electronic Engineering, Soongsil University.