

모션 벡터의 각도 성분 추정을 통한 카메라 움직임 검출

김재호[†], 이장훈^{**}, 장소은^{***}

Camera Motion Detection Using Estimation of Motion Vector's Angle

Jae Ho Kim[†], Jang Hoon Lee^{**}, Soeun Jang^{***}

ABSTRACT

In this paper, we propose a new algorithm that is robust against the effects of objects that are relatively unaffected by camera motion and can accurately detect camera motion even in high resolution images. First, for more accurate camera motion detection, a global motion filter based on entropy of a motion vector is used to distinguish the background and the object. A block matching algorithm is used to find exact motion vectors. In addition, a matched filter with the angle of the ideal motion vector of each block is used. Motion vectors including 4 kinds of diagonal direction, zoom in, and zoom out are added additionally. The experiment shows that the precision, recall, and accuracy of camera motion detection compared to the recent results is improved by 12.5%, 8.6% and 9.5%, respectively.

Key words: Camera Motion Detection, Block Matching Algorithm, Motion Vector, Matched Filter

1. 서 론

최근 인터넷과 스마트폰의 보급화로 인해 많은 사람들이 다양한 영상들을 쉽고 빠르게 접할 수 있게 되었다. 인터넷 TV, 인터넷 방송 등의 실시간 스트리밍 서비스의 수요와 공급이 증가하였으며 유튜브와 페이스북, 트위터와 같은 SNS의 사용으로 개인이 촬영한 영상을 쉽게 공유할 수 있게 되었다. 이로 인해 영상의 검색, 전송, 재생 등의 횟수가 증대되었으며 이러한 현상은 시간이 지날수록 더욱 증가할 것이다. 이렇듯 멀티미디어 시장의 확대와 더불어 촬영 및 편집 기술이 발전하면서 영상 분야의 규모 또한 커지

고 있으며 관련 연구 또한 활발해질 전망이다[1-3].

영화 및 애니메이션 제작에 있어서 카메라 움직임은 흥미를 증진시키는 필수적 요소이며 촬영 기법에 따라 제작자의 의도를 시청자들에게 효과적으로 전달할 수 있게 한다[4]. 특히, 애니메이션에서 캐릭터의 이동 등 동적 구도는 비주얼스토리의 전달력을 높이는 역할을 한다[5-7]. 따라서 카메라 모션을 이해하는 것은 애니메이션이나 영화의 이해에 있어서 매우 중요하고 필수적이다. 뿐만 아니라 CCTV 등 여러 방면에서 카메라 움직임을 검출하여 활용하는 방안이 많이 연구되고 있다[8-10]. 모션 벡터를 통하여 움직임을 검출하는 방식은 예측부호화를 통한 영

※ Corresponding Author : Soeun Jang, Address: (48060) A-2802, Centum3-ro 32, Haeundae-gu, Busan, Korea, TEL : +82-10-2580-3753, FAX : +82-51-743-1165, E-mail : pianee@naver.com

Receipt date : Jun. 14, 2018, Revision date : Jul. 12, 2018
Approval date : Jul. 18, 2018

[†] Dept. of Electronic Eng., Graduate School, Pusan National University (E-mail : jhkim@pusan.ac.kr)

^{**} Dept. of Electronic Eng., Graduate School, Pusan National University
(E-mail : winerquick@naver.com)

^{***} Research Center of QTT, Inc.

※ This work was supported by BK21PLUS, Creative Human Resource Development Program for IT Convergence.

상 압축 방법에 널리 이용되는데, 예를 들어 영상 부호화 방식인 MPEG4, H264, H265 등이 대표적이다. 카메라의 움직임 검출은 동영상압축에 있어서 모션 비트(motion bit) 데이터를 줄이는 데도 중요한 역할을 한다[11].

카메라 움직임 검출을 위한 기존의 연구 방법으로는 SIFT 특징점들을 이용하거나 RANSAC을 이용하여 모션 벡터를 검출한 방법이 있다[12-13]. 또 움직임 추정을 위한 방법으로 옵티컬 플로우(optical flow)를 사용하기도 한다[14]. 하지만 SIFT와 RANSAC을 이용하여 생성한 모션 벡터는 카메라 움직임의 영향을 상대적으로 덜 받는 동적인 물체나 인물에 의해 생성될 확률이 높아서 카메라 움직임 이외의 영향을 받을 확률이 높다.

본 논문에서는 위에서 언급한 기존의 문제점을 해결하고 또 고해상도의 영상에서의 연구와 활용을 위해 보다 정확한 카메라 움직임 검출을 목표로 한다. 영상에서 배경에 해당하는 부분에서만 데이터를 추출하기 위해 블록 정합 알고리즘(block matching algorithm, BMA)[15]을 사용하여 모션 벡터를 생성한다. 글로벌 모션 필터[16]를 적용하여 배경(background)과 객체(object)에 해당하는 부분을 구별한다. 고해상도의 영상에서 추출한 데이터들을 활용하기 위해 Pan과 Tilt의 방향성을 합쳐 대각선 4방향에 대한 카메라 움직임을 추가로 제안한다[17]. 블록 정합 알고리즘으로 추출한 모션 벡터는 블록 별로 이상적인 각도와 정합 필터링(matched filtering)을 통해 구별된다. 각각의 카메라 움직임에 따른 매칭 오차를 계산하여 최소 오차를 통해 최종 카메라 움직임을 검출해낸다.

본 논문에서 제안하는 방법을 검증하기 위한 실험 데이터로 영화 애니메이션을 사용하였다. 애니메이션은 특성상 현실과 유사한 공간과 인물을 인위적으로 만들어 사용하므로 실제 촬영 영상과는 다르게 날씨와 같은 외부 영향이나 카메라 촬영에 있어서의 물리적인 제약이 거의 없이 오직 제작자의 의도에 맞게 만들어지므로 실험 데이터로 사용하기에 적합하다. 그 중에서 아카데미 장편 애니메이션 작품상에 채택 된 작품들 중 ‘겨울왕국(2013)’과 ‘빅 히어로(2014)’를 선정하여 실험을 진행한다.

본 논문은 다음과 같이 구성 된다. 2장에서는 분석 도구들인 모션 벡터(motion vector), 블록 정합 알고

리즘(block matching algorithm, BMA), 글로벌 모션 필터(global motion filter), 정합 필터(matched filter)에 대해 설명한다. 3장에서는 이론적 배경을 바탕으로 배경과 객체의 분리를 통한 배경 모션 벡터 추출, 카메라 움직임 종류 및 정합 필터의 이상적인 벡터각도 생성 및 매칭을 통한 실제 카메라 움직임을 검출 하는 과정을 제안한다. 4장에서는 제안하는 방식을 검증하기 위한 실험 데이터 선정, 다른 알고리즘과의 성능 비교 및 고찰 등을 정리한다. 마지막으로 5장에서는 앞선 실험 결과를 바탕으로 결론을 내린다.

2. 분석 도구

2.1 모션 벡터

MPEG 스트림에서 움직임 보상(motion compensation)을 위해 P-프레임 또는 B-프레임의 매크로 블록 간에 하나 또는 그 이상의 움직임 벡터가 존재 하는데, 이는 움직임 벡터 필드(motion vector field, MVF)로 불려진다[18]. 이 MVF는 다양한 옵티컬 플로우 방식을 이용하여 만들어 질 수 있다. Pan, Tilt, Zoom 등의 방향에 따른 6가지의 카메라 움직임에 대한 모션 벡터를 Fig. 1에 나타내었다. 연속된 두 프레임 간의 MVF는 객체의 독립적인 움직임을 제외 할 경우 카메라의 움직임 방향과 반대 방향의 모션 벡터로 구성된다.

2.2 블록 정합 알고리즘

움직임 추정을 통해 MVF를 만드는 방식인 블록 정합 기법(block matching algorithm, BMA)은 현재

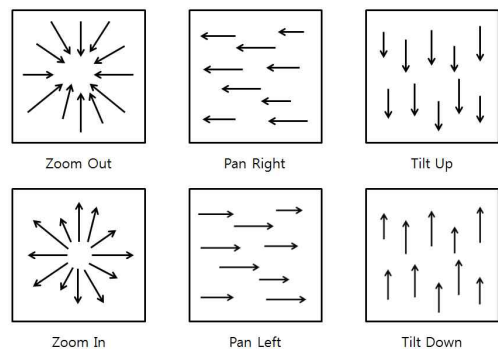


Fig. 1. Direction of motion vector in accordance with camera motion.

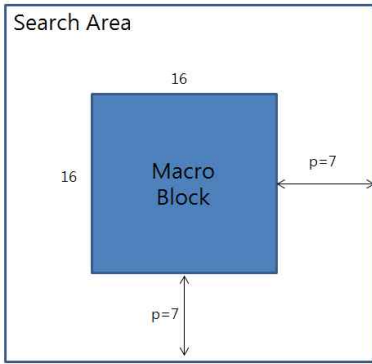


Fig. 2. Block Matching: Macro block of side 16 pixels, search parameter p of size 7 pixels.

프레임의 탐색 범위(search area) 내에서 이전 프레임의 지정된 매크로 블록(macro block)과 가장 유사한 블록을 찾아내는 방법이다[15]. 탐색 범위 내의 여러 개의 후보 블록들 중에서 블록 정합 판단 기준에 따라 가장 작은 비용함수를 가지는 블록의 움직임을 벡터로 정한다. 매크로 블록의 크기가 N , 탐색범위가 p 일 경우, Fig. 2에서와 같이 $N+2p$ 크기의 범위에서 탐색을 실행하게 된다.

블록 정합 기법에 쓰이는 비용함수로는 절대값의 합(sum of absolute differences, SAD), 평균절대오차(mean absolute difference, MAD), 평균제곱오차(mean squared difference, MSD), 상호상관함수(cross correlation function, CCF) 등이 있다[19]. 이 중에서도 계산량에 비해 비교적 정확한 블록 정합결과를 획득 할 수 있는 SAD 방법이 주로 쓰이고 있으며 그 비용함수는 아래와 같이 표현된다.

$$SAD(m,n) = \sum_{a=1}^N \sum_{b=1}^N |x_{a,b} - y_{m+a,n+b}| \quad (1)$$

여기서 m 과 n 은 각각 현재 프레임에서 찾고자하는 블록의 위치와 과거 프레임에서 기준블록의 위치의 차이인 변위, 즉 모션 벡터를 나타내고, 그 블록에서의 픽셀 값을 각각 x 와 y 로 나타내었다. 한편, 블록 정합 알고리즘으로 만들어진 모션 벡터는 프레임에 따라 물체의 크기가 급격하게 변하거나 블록간의 차이가 거의 없을 경우 정확한 모션 벡터를 계산하기 어렵다는 단점이 있다.

2.3 글로벌 모션 필터

영상 검색에 영상의 움직임을 활용하기 위한 방안

으로 글로벌 모션 필터가 제안된 바 있다[16]. 글로벌 모션 필터는 모션 벡터의 움직임의 크기와 방향의 엔트로피를 이용하여 영상의 특징적인 이미지를 만드는 역할을 한다. 본 논문에서도 영상의 배경과 객체를 분리하는 방법으로 글로벌 모션 필터를 적용하였다.

카메라가 움직이면서 촬영한 영상의 경우 배경에 해당하는 영역은 카메라의 움직임에 따라 종속적으로 변한다. 반면에 객체 영역은 주로 영상에서 집중해야하는 부분이므로 인물의 팔, 다리 등의 움직임과 같이 무작위로 움직이며 카메라의 움직임에 크게 영향을 받지 않는다. 따라서 객체영역을 제외한 부분의 영상으로 카메라 움직임을 판별하는 것이 더 정확하기 때문에 글로벌 모션 필터를 적용한다.

연속된 두 프레임 $f, f+1$ 에서 블록 정합 알고리즘으로 추출한 (i,j) 블록의 x 축 방향의 크기는 $\Delta u_{i,j}(f,f+1)$ 이며 y 축 방향의 크기는 $\Delta v_{i,j}(f,f+1)$ 이다. 이때의 모션 벡터의 크기 $M_{i,j}(f,f+1)$ 와 각도 $\theta_{i,j}(f,f+1)$ 는 다음과 같이 표현 할 수 있다.

$$M_{i,j}(f,f+1) = \sqrt{\Delta u_{i,j}^2(f,f+1) + \Delta v_{i,j}^2(f,f+1)} \quad (2)$$

$$\theta_{i,j}(f,f+1) = \tan^{-1} \frac{\Delta v_{i,j}(f,f+1)}{\Delta u_{i,j}(f,f+1)},$$

$$\theta_{i,j}(f,f+1) \in [0, 2\pi) \quad (3)$$

식(2)에서 구한 모션 벡터들의 크기는 블록의 크기에 따라 정해진 m_{th} 로 정규화 과정을 거친다.

$$\hat{M}_{i,j}(f,f+1) = \begin{cases} \frac{M_{i,j}(f,f+1)}{m_{th}}, & \text{if } M_{i,j}(f,f+1) < m_{th} \\ 1, & \text{otherwise} \end{cases} \quad (4)$$

정규화된 모션 벡터 크기의 평균을 Fig. 3에서 보듯이 (i,j) 블록 모션 에너지 $BME_{i,j}$ 로 나타낸다.

$$BME_{i,j} = \frac{1}{N} \sum_{f=0}^{N-1} \hat{M}_{i,j}(f,f+1), \quad BME_{i,j} \in [0, 1] \quad (5)$$

전체 프레임 $f = 0, 1, \dots, N-1$ 에서 카메라의 움직임이 일정하다고 가정하였을 때 (i,j) 블록의 방향은 일정해야 한다. 각도에 대한 엔트로피는 높은 값을 가질수록 방향이 일정하지 않고 낮을수록 방향이 일정할 것이다. 모션 벡터의 각도를 r 등분하여 k 번째 히스토그램인 $AH_{i,j}(k)$ 를 구한 뒤 확률 분포 함수인 $d_{i,j}(p)$ 를 계산한다.

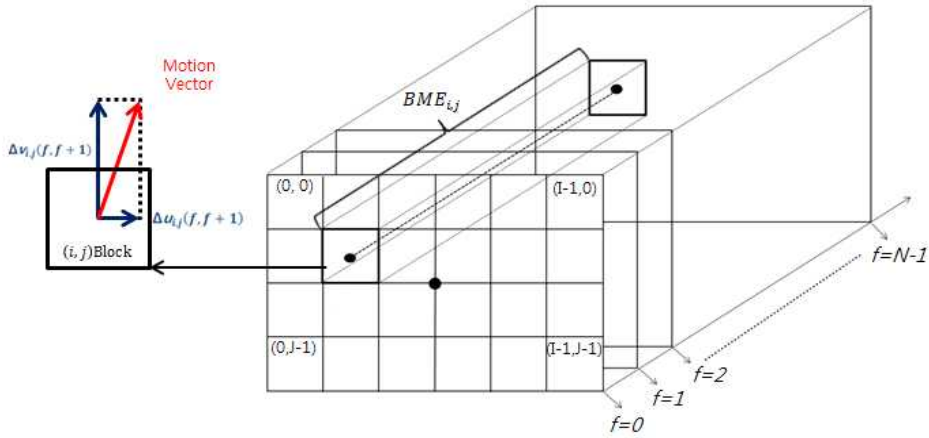


Fig. 3. (i, j) Block and $BME_{i,j}$ in global motion filter.

$$d_{i,j}(k) = \frac{AH_{i,j}(k)}{\sum_{p=1}^r AH_{i,j}(p)} \quad (6)$$

식 (6)를 이용하여 표현한 각도의 엔트로피 $AngEnp_{i,j}$ 는 다음과 같다.

$$AngEnp_{i,j} = - \sum_{p=1}^r d_{i,j}(p) \log d_{i,j}(p), \quad AngEnp_{i,j} \in (0, \log r) \quad (7)$$

식 (7)에서의 $AngEnp_{i,j}$ 를 정규화하여 표현한 객체의 움직임 비율 $RatioBOM_{i,j}$ 은 다음과 같다.

$$RatioBOM_{i,j} = \frac{AngEnp_{i,j}}{\log r}, \quad RatioBOM_{i,j} \in (0, 1] \quad (8)$$

$RatioBOM_{i,j} \in (0, 1]$ 에서 $RatioBOM_{i,j}$ 는 0에 가까울수록 일정한 방향을 가지며 1에 가까울수록 무작위의 방향성을 가진다는 것을 의미한다. 식(5)와 식(8)을 이용한 블록 객체 모션 에너지 $BOME_{i,j}$ 는 다음과 같이 나타낼 수 있다.

$$BOME_{i,j} = BME_{i,j} \times RatioBOM_{i,j} \quad (9)$$

즉, $BOME_{i,j}$ 의 값이 0에 가까울수록 (i, j) 블록은 카메라의 움직임에 근접한 블록이며 1에 가까울수록 카메라의 움직임과는 무관한 블록이 된다.

2.4 정합 필터

정합 필터는 특정 입력 신호에 대하여 신호대 잡음비(signal-to-noise ratio, SNR)가 최대의 출력값을 가지는 선형 시불변 시스템 중의 하나이다. 입력 $i(x, y)$ 에서 신호성분을 $s(x, y)$ 라 하고 잡음이 AWGN

(additive white gaussian noise)라 가정하면 임펄스 응답 함수 $h(x, y)$ 는 다음과 같이 정의 된다[20-22].

$$h(x, y) = Ks(x_0 - x, y_0 - y) \quad (10)$$

여기서, $x = x_0, y = y_0$ 일 때 SNR 출력이 최대가 되며, $K = 2/N_0$ 가 된다. 그 때의 출력 $z(x, y)$ 는 아래와 같다.

$$z(x, y) = \frac{2}{N_0} i(x, y) * s(x_0 - x, y_0 - y) \quad (11)$$

원신호의 종류가 N 개인 경우 신호 성분을 $S_l(x, y)$, $l = 1, \dots, N$ 으로 표현하고, 식(10)과 동일하게 $h_l(x, y)$ 를 여러개 필터로 사용 하여 Fig. 4와 같이 정합필터 시스템을 구성 할 수 있다. 각 필터의 출력 z_1, \dots, z_N 에서 최대 출력인 k 인덱스 값으로 $i(x, y)$ 의 원신호가 $s_k(x, y)$ 이었음을 나타낸다.

3. 제안하는 방법

3.1 카메라 움직임 종류 및 배경 모션 벡터 추출

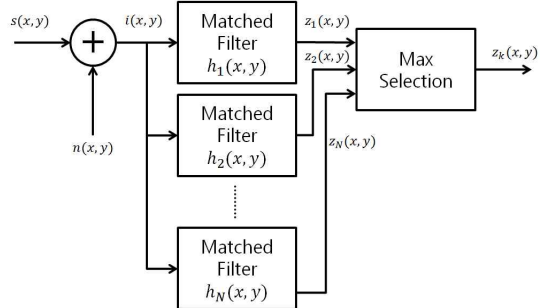


Fig. 4. Block diagram of matched filter.

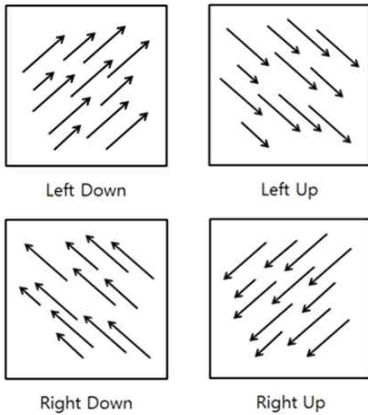


Fig. 5. Movement of motion vector in 4 kinds of diagonal direction.

고해상도 영상의 경우 저해상도보다 더 많은 픽셀을 사용하여 영상을 만들기 때문에 모션 벡터 또한 더 많이 추출해 낼 수 있다. 결과적으로 활용 할 수 있는 정보 또한 많아진다. 이를 활용하는 방안으로 Pan과 Tilt의 방향성을 합하여 대각선 4방향인 Left Down (LD), Left Up (LU), Right Down (RD), Right Up (RU)을 Fig. 5에서처럼 추가하여 분류 하였다.

기존 6가지 움직임과 마찬가지로 대각선 4가지 방향의 모션 벡터 또한 카메라의 움직임과 반대 방향으로 생성 된다. 본 논문에서 사용하는 카메라 움직임 분류인 MT를 다음과 같이 10가지로 정의한다.

$$MT \in \{ \text{Pan Left, Pan Right, Tilt Up, Tilt Down, Zoom In, Zoom Out, LD, LU, RD, RU} \} \quad (12)$$

한편, 글로벌 모션 필터를 이용하여 배경 모션 벡터를 추출하고자 한다. 먼저 각 블록에 해당하는 $BOME_{i,j}$ 를 구한다. 배경과 객체의 구별하는 문턱값(threshold)으로 Th_{BOME} 를 사용하며 이는 $BOME_{i,j}$ 의 평균으로 한다.

$$Th_{BOME} = \frac{1}{I \times J} \sum_{j=0}^{J-1} \sum_{i=0}^{I-1} BOME_{i,j} \quad (13)$$

$BOME_{i,j}$ 가 Th_{BOME} 보다 작은 값을 가지는 블록은 방향성이 일정한 배경에 해당하는 블록으로 간주하고 Th_{BOME} 보다 높을 경우 방향성이 일정하지 않은 객체에 해당하는 블록으로 간주한다.

3.2 정합 필터의 이상적인 벡터 각도 생성 및 오차 비교

블록에서 생성되는 각각의 모션 벡터는 카메라 움직임과 프레임에서 블록의 위치에 따라 달라질 것이다. 뿐만 아니라 영상의 해상도에 따라 모션 벡터가 달라 질 것이므로 이러한 변수들을 고려하여 3.1절에서 제시한 카메라 움직임 타입에 따른 이상적인 벡터 각도, $\phi_{i,j}^{MT}$ 를 제안한다. 프레임의 너비, 높이를 각각 W, H 로 두고, 벡터의 시작점을 $a = i - W/2, b = -j + H/2$ 라고 각각 정의한다. 식 (14)에서 (21)까지는 각각 Pan Left, Pan Right, Tilt Down, Tilt Up, LD, LU, RD, RU에 대한 $\phi_{i,j}^{MT}$ 이며 블록의 위치에는 관계없는 일정한 이상적인 각도이다.

$$\phi_{i,j}^{Pan\ Left} = 0^\circ \quad (14)$$

$$\phi_{i,j}^{Pan\ Right} = 180^\circ \quad (15)$$

$$\phi_{i,j}^{Tilt\ Down} = 90^\circ \quad (16)$$

$$\phi_{i,j}^{Tilt\ Up} = 270^\circ \quad (17)$$

$$\phi_{i,j}^{LD} = 45^\circ \quad (18)$$

$$\phi_{i,j}^{LU} = 315^\circ \quad (19)$$

$$\phi_{i,j}^{RD} = 135^\circ \quad (20)$$

$$\phi_{i,j}^{RU} = 225^\circ \quad (21)$$

한편, 식 (22)과 (23)는 각각 Zoom In과 Zoom Out에 대한 $\phi_{i,j}^{MT}$ 이며 Fig. 6에서 보듯이 블록의 위치에

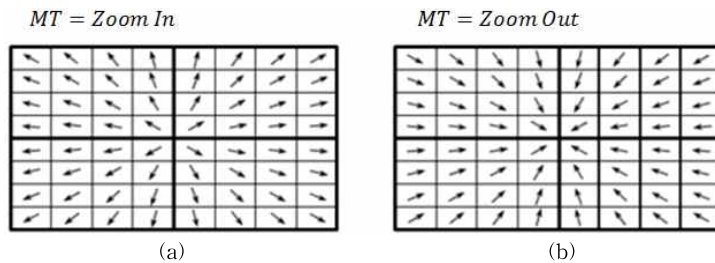


Fig. 6. Ideal angle of Zoom In and Zoom Out, (a) $\phi_{i,j}^{Zoom\ In}$ and (b) $\phi_{i,j}^{Zoom\ Out}$.

따라 각각의 방향이 다르다.

$$\Phi_{i,j}^{Zoom In} = \begin{cases} \frac{180}{\pi} \tan^{-1}\left(\frac{b}{a}\right)^\circ, & a \geq 0, b \geq 0 \\ \left(\frac{180}{\pi} \tan^{-1}\left(\frac{b}{a}\right) + 180\right)^\circ, & a < 0, b \geq 0 \\ \left(\frac{180}{\pi} \tan^{-1}\left(\frac{b}{a}\right) + 180\right)^\circ, & a < 0, b < 0 \\ \left(\frac{180}{\pi} \tan^{-1}\left(\frac{b}{a}\right) + 360\right)^\circ, & a \geq 0, b < 0 \end{cases} \quad (22)$$

$$\Phi_{i,j}^{Zoom Out} = \begin{cases} \left(\frac{180}{\pi} \tan^{-1}\left(\frac{b}{a}\right) + 180\right)^\circ, & a \geq 0, b \geq 0 \\ \left(\frac{180}{\pi} \tan^{-1}\left(\frac{b}{a}\right) + 360\right)^\circ, & a < 0, b \geq 0 \\ \frac{180}{\pi} \tan^{-1}\left(\frac{b}{a}\right)^\circ, & a < 0, b < 0 \\ \left(\frac{180}{\pi} \tan^{-1}\left(\frac{b}{a}\right) + 180\right)^\circ, & a \geq 0, b < 0 \end{cases} \quad (23)$$

추정된 모션벡터의 각도 $\theta_{i,j}(f, f+1)$ 와 이상적인 벡터 각도 $\Phi_{i,j}^{MT}$ 는 항상 양수값을 가진다. 매칭의 과정으로 Th_{BOME} 보다 작은 부분의 블록의 모션 벡터와 앞서 $\Phi_{i,j}^{MT}$ 와 비교하여 각도 차이를 나타내는 $AngDiff_{i,j}^{MT}(f, f+1)$ 을 식(23)을 통해 구한다. 다음 과정으로 카메라 움직임 타입에 따른 오차 에너지 $e_{i,j}^{MT}(f, f+1)$ 를 계산한다.

$$AngDiff_{i,j}^{MT}(f, f+1) = \left| \frac{180(\theta_{i,j}(f, f+1))}{\pi} - \Phi_{i,j}^{MT} \right|, \quad (23)$$

$$AngDiff_{i,j}^{MT}(f, f+1) \in [0, 180]$$

$$e_{i,j}^{MT}(f, f+1) = \left(\frac{AngDiff_{i,j}^{MT}(f, f+1)}{180} \right)^2, \quad (24)$$

$$e_{i,j}^{MT}(f, f+1) \in [0, 1]$$



(a)



(b)

Fig. 7. Animations used in experiment, (a) Frozen and (b) Big Hero.

여기서, $e_{i,j}^{MT}(f, f+1)$ 은 실제 벡터 $\theta_{i,j}(f, f+1)$ 가 $\Phi_{i,j}^{MT}(i, j)$ 와 얼마나 차이가 있는가를 나타낸다. $e_{i,j}^{MT}(f, f+1)$ 의 값은 0에 가까울수록 모션 벡터의 각도와 이상적인 벡터각도의 방향이 비슷하며, 1에 가까울수록 서로 정반대의 방향을 가지고 있다는 것을 나타내며, 이때의 오차 에너지의 합인 E^{MT} 은 다음과 같다.

$$E^{MT} = \sum_{f=0}^{N-1} \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} e_{i,j}^{MT}(f, f+1) \quad (23)$$

계산된 총 10개의 E^{MT} 중에 가장 작은 값을 가지는 MT 를 카메라 움직임 타입으로 결정한다.

4. 실험 결과 및 고찰

4.1 실험 대상 선정

본 논문에서 제안하는 방법을 검증하기 위한 실험 데이터로 영화 애니메이션을 사용하였다. 애니메이션은 특성상 현실과 유사한 공간과 인물을 인위적으로 만들어 사용하고 있으며, 이는 실제 촬영 영상과는 다르게 날씨와 같은 외부 영향이나 카메라 촬영에 있어서의 물리적인 제약이 거의 없이 제작자의 의도에 맞게 만들어지므로 실험 데이터로 사용하기에 적합하다.

영화 애니메이션 중에서 아카데미 장편 애니메이션 작품상에 채택된 작품들 중 두 편을 선택하였다. 2013년에 개봉한 ‘겨울왕국’과 2014년 개봉작인 ‘빅 히어로’에서 실험을 위한 영상 데이터를 추출하여 실험을 진행하였다.

애니메이션의 특성상 현실에서 나타나기 힘든 폭발 및 마법과 같은 특수효과가 화면 전체에 영향을 미치는 부분의 영상은 실험 데이터에서 제외하였다. 또한 비나 눈과 같은 날씨 등의 영향으로 카메라 움직임을 제대로 분간 할 수 없는 경우의 영상 또한 실험 데이터에서 제외 하였다. 최종적으로 추출된 테스트용 영상은 ‘겨울 왕국’에서 335개, ‘빅 히어로’에서 210개 이다. ‘겨울 왕국’에서 추출한 데이터의 총 길이는 25분 40초로 전체 길이의 23.7%에 해당하며, ‘빅 히어로’에서는 17분 6초로 전체 길이의 15.8%에 해당한다.

4.2 알고리즘 실험 결과 및 비교 분석

본 논문에서 제시하는 방법의 성능을 검증하기 위

Table 1. Comparison of Precision, Recall and Accuracy

	Precision		Recall		Accuracy	
	PM	Ma	PM	Ma	PM	Ma
Pan	0.9403	0.8275	0.8869	0.8674	0.9666	0.8761
Tilt	0.9179	0.7250	0.7951	0.6904	0.9714	0.8857
Zoom	0.9583	0.8888	0.8868	0.7529	0.9713	0.8619
Average	0.9388	0.8137	0.8562	0.7702	0.9697	0.8745

해 다른 연구자가 제시한 카메라 움직임 검출 알고리즘과 비교하였다. TRECVID에서 제시하는 기준과 비교할 때 Pan과 Tilt에서 기존의 측정된 Recall 보다 더 높은 수치를 보인 [12]의 알고리즘을 본 실험의 비교 대상으로 선택하였다. 실험 데이터로 아카데미 장편 애니메이션 작품상 수상작인 ‘빅 히어로’에서 추출한 210개의 샷(shot)을 사용했다. Pan, Tilt, Zoom 각각에 대하여 본 논문에서 제안하는 방식(PM)과 SIFT 알고리즘[23]을 이용한 [12]의 방식에 대한 Precision, Recall, Accuracy [24]값을 아래 Table 1에 정리하였다.

Precision은 평균 12%, Recall의 경우 7%, Accuracy는 9% 이상 높아진 정확도를 확인 할 수 있었다. 즉, 모든 측면에서 제안한 방법의 정확도가 비교대상 방식보다 높은 결과 값을 보였다.

한편, 제안하는 알고리즘을 이용하여 추정한 카메라 움직임에 대한 결과를 살펴보자. 앞 4.2절에서의 결과를 바탕으로 하여 제안하는 방식을 2013년 개봉작인 ‘겨울 왕국’에서 추출한 335개의 데이터와 ‘빅 히어로’에서 추출한 210개의 데이터에 적용시켜 실험 하였으며 아래 Table 2에서 그에 대한 결과를 보인다.

Fig. 8은 ‘겨울 왕국’에서 추출한 실험 데이터이며 Zoom In의 카메라 움직임을 본 논문에서 제안하는 방식으로 올바르게 찾아낸 실험 데이터 영상의 예시이다. 샷 안에서의 카메라 움직임이 한가지이며 객체

에 해당하는 인물의 뒷모습을 제외한 부분인 배경에 해당하는 영역이 넓다. 이 영상에서 각 MT 에 대한 E^{MT} 은 다음과 같다. 특히, Zoom In에서의 오차 에너지가 낮은 것을 알 수 있다. 그러나, Zoom의 중심점이 화면의 중앙이 아닌 경우 등에서 카메라 움직임을 검출하기 힘들었다.

Fig. 9의 그림은 ‘빅 히어로’에서 추출한 실험 데이터이며 Pan Right의 카메라 움직임을 본 논문에서 제안하는 방식으로 올바르게 찾아낸 실험 데이터 영상의 예시이다. 시간이 흐름에 따라 Pan Right의 카메라 움직임과 함께 화면내의 인물들과 애완동물이

Table 2. Precision, Recall and Accuracy of Proposed Method

	Precision	Recall	Accuracy
Pan Left	0.9750	0.7800	0.9642
Pan Right	0.9286	0.8125	0.9642
Tilt Up	1.0000	0.8000	0.9851
Tilt Down	0.7241	0.9130	0.9701
Zoom In	1.0000	0.8036	0.9672
Zoom Out	0.9020	0.9020	0.9701
LD	0.5714	0.9524	0.9522
LU	0.7667	0.9200	0.9731
RD	0.9500	0.9048	0.9910
RU	0.6522	1.0000	0.9745
Average	0.8470	0.8788	0.9711

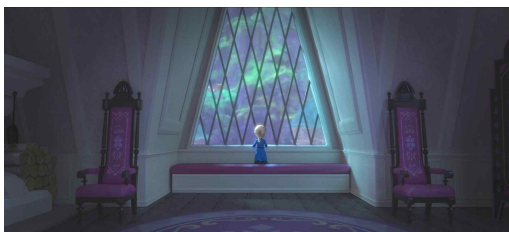


Fig. 8. Example of experiment data (Zoom In).

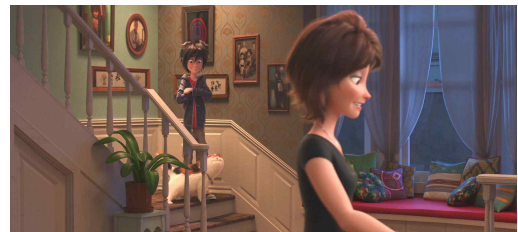


Fig. 9. Example of experiment data (Pan Right).

Table 3. E^{MT} on each MT in Fig. 8.

Camera Motion	Pan Left	Pan Right	Tilt Up	Tilt Down	Zoom In	Zoom Out	LD	LU	RD	RU
E^{MT}	1228.01	1516.95	2491.95	746.95	447.80	2328.58	931.25	1600.18	904.65	2030.21

Table 4. E^{MT} on each MT in Fig. 9.

Camera Motion	Pan Left	Pan Right	Tilt Up	Tilt Down	Zoom In	Zoom Out	LD	LU	RD	RU
E^{MT}	31742.91	13996.80	20135.88	21703.61	21461.34	23145.22	30481.20	26021.73	16528.52	15827.43

Table 5. E^{MT} on each MT in Fig. 10.

Camera Motion	Pan Left	Pan Right	Tilt Up	Tilt Down	Zoom In	Zoom Out	LD	LU	RD	RU
E^{MT}	5931.98	5849.16	6511.09	5813.84	6290.90	5589.69	6015.16	6457.22	6255.14	6538.15

움직이며 이외의 실내 배경은 카메라의 움직임과 반대인 왼쪽으로 일정하게 이동하는 것을 관찰 할 수 있다. 이 예에서도 역시 Pan Right의 오차 에너지가 가장 적음을 알 수 있다.

한편, 실험 결과에서 검출 오류에 대한 고찰을 해 보고자 한다. Fig. 10은 Tilt Down의 카메라 움직임을 본 논문에서 제안하는 방식으로 잘못 찾아낸 실험 데이터 영상의 예시이다. 실제 결과 값은 Tilt Down이지만 추정 결과는 Zoom Out로 측정 되었다. 그 이유는 인물의 움직임에 배경의 대부분이 가려지며 눈이 화면의 대부분을 차지하고 있기 때문에 유효한 모션 벡터를 제대로 찾지 못했기 때문이다.

본 실험에서 검출된 오류의 가장 큰 이유는 모션 벡터의 생성에 결과가 많은 영향을 받는 것이다. 블록 정합 알고리즘으로 추출하는 모션 벡터가 정확하고 많은 수의 벡터를 생성한다면 본 논문에서 제안하는 알고리즘으로 정확한 카메라 움직임을 추출해 낼 수 있지만 화면의 전반적인 밝기와 색상의 단조로움



Fig. 10. Example of experiment data (Tilt Down).

에 따라 모션 벡터의 생성이 어려울 경우 정확성이 감소한다. 또한 객체의 움직임이 화면의 대부분의 차지 할 경우 카메라의 움직임과 관련이 있는 배경 모션 벡터의 양이 현저히 적어지기 때문에 알고리즘의 정확성이 떨어진다. 이와 같은 단점을 보완하기 위한 방향으로 모션 벡터를 더 정확하게 추출해 낼 수 있는 알고리즘을 적용하면서 스토리 전개에 있어서 더 세분화된 단계[25]에 적용한다면 더 나은 결과를 낼 수 있을 것이다.

5. 결 론

본 논문에서는 비교적 카메라의 움직임에 영향을 덜 받는 객체의 영향에 강인할 뿐만 아니라 고해상도 영상에서도 카메라 움직임을 정확하게 검출해 낼 수 있는 새로운 알고리즘을 제안하였다. 먼저, 보다 정확한 카메라 움직임 검출을 위하여 모션 벡터의 엔트로피를 기반으로 한 글로벌 모션 필터를 사용하여 배경과 객체를 구분하였고, 블록 정합 알고리즘을 이용하여 고해상도영상을 적절한 크기의 매크로블록으로 나누어 모션벡터를 찾아내었다. 또, 각 블록의 이상적인 모션 벡터의 각도와와의 정합필터를 통해 카메라 움직임을 검출하였다. 제안된 방식을 통한 카메라 움직임 검출의 결과가 비교대상 알고리즘 보다 Precision은 평균 12.5%, Recall의 경우 8.6%, Accuracy는 9.5% 만큼 개선되었다.

실험 데이터를 통한 분석 결과를 보면 대부분 정

확도가 높았으나, 화면의 밝기가 어둡거나 배경이 단색 위주로 이루어져 모션 벡터를 제대로 찾을 수 없는 경우, 또 Zoom의 중심점이 화면의 중앙이 아닌 경우 등에서 카메라 움직임을 검출하기 힘들었다. 향후 이러한 문제점들에 강인한 새로운 연구가 필요할 것이며, 또 애니메이션뿐만 아니라 더 다양한 배경을 보여주는 영화에도 적용 가능할 것이다.

REFERENCE

- [1] A. Hampapur, R. Jain, and T.E. Weymouth, "Production Model Based Digital Video Segmentation," *Multimedia Tools and Applications*, Vol. 1, No. 1, pp. 9-46, 1995.
- [2] Y.F. Ma, L. Lu, H.J. Zhang, and M. Li, "A User Attention Model for Video Summarization," *Proceedings of 10th ACM International Conference on Multimedia*, pp. 533-542, 2002.
- [3] Y. Li, S.H. Lee, C.H. Yeh, and C.J. Kuo, "Techniques for Movie Content Analysis and Skimming: Tutorial and Overview on Video Abstraction Techniques," *IEEE Signal Processing Magazine*, Vol. 23, No. 2, pp. 79-89, 2006.
- [4] Y.J. Lee and H.T. Kim, "Analysis of Movie Production by Camera Movement," *Korea Digital Design Council*, Vol. 12, No. 1, pp. 491-501, 2012.
- [5] S. Jang, Y.S. Kim, and J.H. Kim, "Fusion Analysis of Visual Story Telling and Original Sound Track of 'Beauty and the Beast' Animation," *Korea Science and Art Forum*, Vol. 23, pp. 295-307, 2016.
- [6] S. Jang, L. Lou, and J.H. Kim, "Integrated Analysis of Visual Story Telling and Original Sound Track of 'Alladin' Animation," *Korea Science and Art Forum*, Vol. 24, pp. 375-388, 2016.
- [7] B.S. Sung and J.H. Kim, "Analysis of Kinetic Composition for the Visual Story Telling of Animation," *Korean Society of Design Science*, Vol. 22, No. 3, pp. 273-284, 2009.
- [8] S. Pumrin and D.J. Dailey, "Roadside Camera Motion Detection for Automated Speed Measurement," *Proceedings of IEEE 5th International Conference on Intelligent Transportation Systems*, pp. 147-151, 2002.
- [9] P. Bouthemy, M. Gelgon, and F. Ganansia, "A Unified Approach to Shot Change Detection and Camera Motion Characterization," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 9, No. 7, pp. 1030-1044, 1999.
- [10] M. Naito, K. Matsumoto, K. Hoashi, and F. Sugaya, "Camera Motion Detection Using Video Mosaicing," *Proceeding of 2006 IEEE International Conference on Multimedia and Expo*, pp. 1741-1744, 2006.
- [11] A. Jain, "Bit Reduction Based Matching Criterion for Motion Compensation in Video Coding," *IEEE International Multimedia, Signal Processing and Communication Technologies*, pp. 36-39, 2009.
- [12] S. Ma and W. Wang, "Effective Camera Motion Analysis Approach," *Proceeding of IEEE International Conference on Networking, Sensing and Control*, pp. 111-116, 2010.
- [13] T. Thorsten, N. Hasler, Michael Wand, H.P. Seidel, "Merging of Feature Tracks for Camera Motion Estimation from Video," *Proceeding of IET 5th European Conference on Visual Media Production*, 2008.
- [14] D. Murray and A. Basu, "Motion Tracking with an Active Camera," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 5, pp. 449-459, 1994.
- [15] A. Barjatya, "Block Matching Algorithms for Motion Estimation," *IEEE Transactions on Evolution Computation*, Vol. 8, No. 3, pp. 225-239, 2004.
- [16] Y.F. Ma and H.J. Zhang, "A New Perceived Motion Based Shot Content Representation," *Proceedings of IEEE International Conference on In Image Processing*, pp. 426-429, 2001.

[17] J.H. Lee, *Camera Motion Detection Using Matched Filter with Motion Vector's Angle*, Master's Thesis of Pusan National University of Technology, 2016.

[18] J. G. Kim, H.S. Chang, J. Kim, H.M. Kim., "Efficient Camera Motion Characterization for MPEG Video Indexing," *Proceeding of IEEE International Conference on Multimedia and Expo*, 2000.

[19] Z.L. He, C.Y. Tsui, K.K. Chan, and M.L. Liou, "Low-power VLSI Design for Motion Estimation Using Adaptive Pixel Truncation," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 10, No. 5, pp. 669-678, 2000.

[20] B. Sklar, *Digital Communications: Fundamentals and Applications*, Prentice Hall, New Jersey, 2001.

[21] G.L. Turin, "An Introduction to Matched Filters," *Proceedings of the IEEE*, Vol. 64, No. 7, pp. 1092-1112, 1976.

[22] G.S. Kim, *The Recognition of Car Plate based on Normalized Matched Filter*, Master's Thesis of Pusan National University, 2005.

[23] D.L. Olson and D. Delen, *Advanced Data Mining Techniques*, Springer Science and Business Media, Berlin, 2008.

[24] D.G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110, 2004.

[25] J.H. Kim, J.H. Lee, and S. Jang, "Animation OST Musical Element Analysis Based on A Narrative Process Classification Model," *Journal of Korea Multimedia Society*, Vol. 17, No. 10, pp. 1239-1252, 2014.



김 재 호

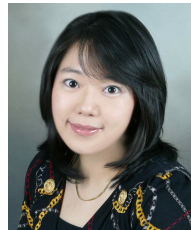
1980년 2월 부산대학교 전기기계공학과 학사
 1982년 2월 한국과학기술원 산업전자공학과 석사
 1990년 2월 한국과학기술원 전기 및 전자공학과 박사

1992년 3월~현재 부산대학교 전자전기공학과 및 영상정보 협동과정 교수
 관심분야 : 애니메이션 캐릭터 자동생성, 영상처리



이 장 훈

2013년 부산대학교 전자전기공학부 공학사
 2015년 부산대학교 전자전기컴퓨터공학과 석사
 관심분야 : 영상처리



장 소 은

1998년 2월 충남대학교 예술대학 음악학과 학사
 2009년 2월 부산대학교 예술대학 대학원 음악학과 석사
 2016년 8월 부산대학교 공과대학 대학원 영상정보협동과정 박사

관심분야 : 영상음악 분석, 음악과 감정의 상관관계, 영상 배경음악