

변형계층적 모션벡터 추정알고리즘

김 정 웅*, 양 해 술*

*호서대학교 벤처전문대학원

e-mail : jwk@korea.com, hsyang@office.hoseo.ac.kr

Metamorphosis Hierarchical Motion Vector Estimation Algorithm

Jeong-Woong Kim*, Hae-Sool Yang*

*Dept. Application of Computer Technology, Hoseo Graduate School of Venture

요 약

다양한 종류의 컴퓨터가 사람, 사물, 환경 속에 내재되어 있고, 이들이 서로 연결되어, 필요한 곳에서 활용할 수 있는 유비쿼터스 환경에서는 홈 네트워크를 통해 이 기종 기기간 다양한 데이터 교환을 요구한다. 더욱이 원활한 영상 데이터의 처리, 전송, 모니터링 기술은 핵심적 요소가 아닐 수 없다. 공간 및 시간적인 해상도, 컬러의 표현 그리고 화질의 측정방법 등 고전적 영상 처리 연구 분야뿐만 아니라 국한된 대역폭을 갖는 홈네트워크의 전송체계에서 전송률 문제에 대한 심도 있는 연구가 필요하다. 본 논문에서는 홈네트워크 상황에서 콘텐츠의 중심이 되는 영상 데이터의 전송과 처리 그리고 제어를 위하여 새로운 움직임 추정 알고리즘을 제안한다. 각도, 거리등 다양한 환경에서 전송되어지는 스테레오 카메라의 영상데이터들은 축소, 확대, 이동, 보정 등 전처리 후 제안된 변형계층 모션벡터 추정 알고리즘을 이용하여 압축 처리, 전송된다. 기존 모션벡터 추정 알고리즘의 장점을 계승하고 단점을 보완한 변형계층 알고리즘은 비정형, 소형 매크로 블록을 이용하여 휘도의 편차가 큰 영상의 효율적 움직임 추정에 이용된다. 본 논문에서 제안한 변형계층 알고리즘과 이를 이용해 구현된 영상시스템은 유비쿼터스 환경에서 다양하게 활용될 수 있다.

1. 서 론

다수의 카메라에서 전송된 다양한 영상을 단안 영상으로 표현 처리하거나 이들을 모니터링 또는 제어하는 것은 홈오토크메이션이나 유비쿼터스 환경에서는 꼭 필요한 요구 기술이다. 그러나 국한된 전송 대역폭을 갖는 홈 네트워크 전송 체계에서 다중카메라로부터 전송된 영상을 축소, 확대, 이동, 회전, 보정 등 여러 전처리 과정을 거쳐 압축, 해독, 전송하여 실시간으로 모니터링, 제어하는 것은 쉬운 문제가 아니다.

본 연구는 한정된 자원을 사용해야 하는 유비쿼터스 환경에서의 핵심 정보인 영상 데이터 처리 방안과 다중 카메라에서 전송되어지는 여러 화면이 단안 화상을 전송하는 것과 비교하여 크게 증가하는 대역폭 문제를 해결하려는데 연구의 목적이 있다.

본 논문에서는 다중 카메라에서 입력된 영상을 기존의 H.264 방식에서 사용하는 계층적인 움직임 벡터 추정 알고리즘을 개선한 변형 알고리즘을 제안한다.

2. 관련연구

2.1. 스테레오 카메라의 상대적인 변화 추정

FFT 알고리즘을 이용하여 주파수 도메인에서 스테레오 카메라가 취득한 영상의 이동, 회전 및 크기의 상대적인 변화를 추정하고, 이를 통해 중첩 영역을 계산해 낸다[1]. 첫 번째, 영상의 이동 변환에 있어서는 수식 1과 같이 푸리에 변환의 시간천이성질을 이용하며, 수식 2와 같이 Cross-power Spectrum 상에서 (x_0, y_0) 에서 임펄스가 발생 했을 때 두 신호가 정합되는 이동 위치를 계산한다. 두 번째, 영상의 회전 변환에서는 극 좌표계를 이용하며, 회전이 있는 두 신호의 관계는 수식 3과 같이 표현할 수 있다. 그리고

※ 본 연구는 정보통신부 지원 ITRC 프로그램의 지원을 받아 수행되었음.

magnitude M_1, M_2 가 있을 때, 위의 성질을 이용하여, 수식 4와 같이 나타낼 수 있다[2].

$$f_2(x, y) = f_1(x - x_0, y - y_0) \xrightarrow{F.T} F_2(u, v) = e^{-j2\pi(ux_0 + vy_0)} F_1(u, v)$$

$$f_2(x, y) = af_1(x - x_0, y - y_0) + b \xrightarrow{F.T} F_2(u, v) = e^{-j2\pi(ux_0 + vy_0)} a \times F_1(u, v) + b \times \delta(u, v) \quad \text{수식(1)}$$

$$\frac{F_1(u, v)F_2^*(u, v)}{|F_1(u, v)F_2^*(u, v)|} = e^{j2\pi(ux_0 + vy_0)} \xrightarrow{I.F.T} \delta(x - x_0, y - y_0) \quad \text{수식(2)}$$

$$f_2(x, y) = f_1(x \cos \theta_0 + y \sin \theta_0, -x \sin \theta_0 + y \cos \theta_0) \xrightarrow{F.T} F_2(u, v) = F_1(u \cos \theta_0 + v \sin \theta_0, -u \sin \theta_0 + v \cos \theta_0) \times e^{-j2\pi(ux_0 + vy_0)} \times a + b \times \delta(u, v) \quad \text{수식(3)}$$

$$M_2(u, v) = |a| M_1(u \cos \theta_0 + v \sin \theta_0, -u \sin \theta_0 + v \cos \theta_0) \quad \text{수식(4)}$$

여기서, 극 좌표계에서 폴라(polar) 좌표계로 변환하게 되면, M_2 는 수식 5와 같이 나타낼 수 있다[9].

$$M_2(\rho, \varphi) = |a| \times M_1(\rho, \varphi - \theta) \quad (\rho = \sqrt{u^2 + v^2}, \varphi = \tan^{-1}(v/u), u = \rho \cos \varphi, v = \rho \sin \varphi) \quad \text{수식(5)}$$

따라서, M_1, M_2 를 H_1, H_2 로 FT를 수행하면 수식 6을 얻을 수 있으며, 이를 이용하여 Cross-power spectrum 상에서 (x_0, y_0) 에서 임펄스가 발생했을 때 두 신호가 정합되는 회전 위치를 계산한다. 마지막으로, 영상의 크기 변환은 수식 7과 같은 푸리에 변환 스케일 성질(scale property)을 이용하며, 다음과 같이 수식 8을 이용하여 스케일 변환을 시키게 된다[10].

$$\frac{H_1(v, \beta)H_2^*(v, \beta)}{|H_1(v, \beta)H_2^*(v, \beta)|} = e^{j2\pi(\beta \times \theta)} \xrightarrow{I.F.T} \delta(0, \theta), \quad H_2(v, \beta) = |a| \times e^{-j2\pi(\beta \times \theta)} H_1(v, \beta) \quad \text{수식(6)}$$

$$f_2(x, y) = f_1(x/a, y/b) \xrightarrow{F.T} F_2(u, v) = \frac{1}{|ab|} F_1(u/a, v/b) \quad \text{수식(7)}$$

$$M_2(\log \rho, \varphi) = M_1(\log \rho - \log a, \varphi - \theta) \quad \rho = \sqrt{u^2 + v^2}, \varphi = \tan^{-1}(v/u), u = \rho \cos \varphi, v = \rho \sin \varphi \quad \text{수식(8)}$$

2.2 중첩 영역의 정보를 이용한 컬러 값 보정

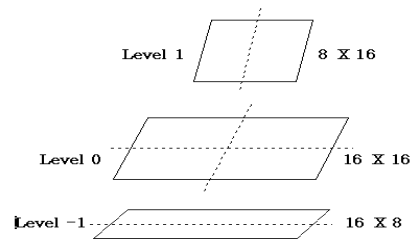
스테레오 영상을 획득하는 과정에서 카메라의 불일치와 외부 환경에 의한 불균형 문제로 발생하는 오류는 전처리 과정에서 수식 9와 같이 컬러 값 보정(Balancing)으로 해결한다.

$$I_R(i, j) = \frac{\sigma_L}{\sigma_R} (I_R(i, j) - \mu_R) + \mu_L, \forall (i, j) \in I_R \quad \text{수식(9)}$$

3. 변형계층 움직임 추정 알고리즘

3.1 계층적 움직임 벡터 추정 알고리즘의 변형

기존의 계층적 방법은 두 가지, 즉 휘도의 평균(mean intensity)을 통한 방법과 부 표본을 추출(subsampling)하는 방법으로 구성하나 변형 방법은 상위 휘도 값과 하위 휘도 값(평균 기준)을 이분적으로 분리하여 저해상도 영상 매크로 블록을 구성한다. (그림 1)에서는 참조 영상에 대해 Level 1과 Level -1에서 이분적으로 분리된 두 장의 저해상도 영상을 보이고 있다. 분명한 비교를 위하여 현재영상에 대해서도 참조영상과 유사하게 저해상도 영상들을 얻어야 한다.

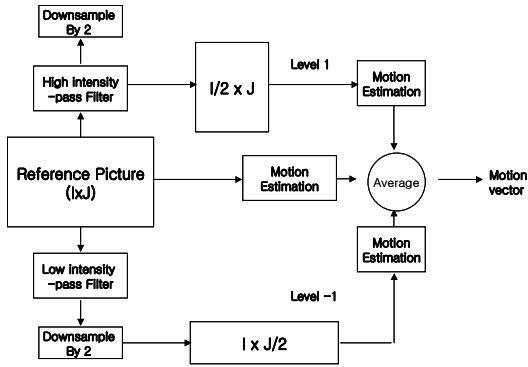


(그림 1) 영상의 휘도에 따른 계층적 구조

그런 다음 이전 이미지에서도 움직임 추정이 세밀하게 이루어져야 한다. 이를 위해 휴리스틱 탐색기법을 적용해서 상, 하 단계에서의 최적의 정합 영역을 찾은 후에 원본 단계로 그 결과를 전파하는 것이다. 변형 계층법은 벡터 값의 평균을 더하는 방법으로 다음 단계에 전파된다.

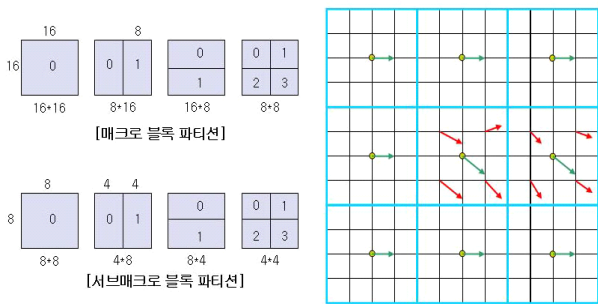
구체적 변형 계층적 움직임 벡터 추정 알고리즘의 내용은 (그림 2)에 잘 표현되어 있다. 먼저, 원본영상 내에 (x, y) 에 위치한 매크로 블록을 가정해보자. 이 매크로 블록의 Level 1과 Level -1에서 해당 매크로 블록은 각각 $(x/2, y)$ 과 $(x, y/2)$ 에 위치해 있다. 그리고 기존과 동일하게 Level 0에 해당하는 매크로 블록을 16×16 탐색 영역을 $[-m, +m]$ 라고 하자. 움직임 벡터의 탐색은 Level 1과 Level -1에서 동시에 시작한다. Level 1단계에서는 매크로 블록이 8×16 크기를 Level -1단계에서는 매크로 블록이 16×8 크기를 가지고 있다. 또한, 탐색 영역에 있어서도 변화가 있다. 즉 각각 $[-m/2, +m/2]$ 탐색 $1/2$ 영역으로 감소하게 된다. 그래서 Level 1에서는 $[-m/2, +m/2]$ 의 탐색 영역과 8×16 매크로 블록에 대해 전역 탐색, 3-Step 탐색법등을 사용할 수 있다. 여기서 8×16 매크로 블록의 원점은 $(x/2, y)$ 이고 탐색 영역은 그 원점을 중심으로 $[-1, +1]$ 화소의 범위이다. 현재 영상과 참조 영상을 비교하여 얻어진 움직임 벡터를 $(i_1$

, j_1)이라고 하자. Level -1에서는 16×8 매크로 블록에 대해 움직임 추정을 수행하게 되는데, 이때 움직임 추정의 원점은 $(x, y/2)$ 이고 탐색영역은 $[-1, +1]$ 화소의 범위이다. 본 단계에서 얻어진 움직임 벡터를 (i_{-1}, j_{-1}) 라고 하자.



(그림 2) 변형 계층 움직임 추정 방법

Level 0에서는 16×16 매크로 블록에 대해 움직임 추정을 수행하게 되는데, 이때 움직임 추정의 원점은 $(x+(i_1+i_{-1})/2, y+(j_1+j_{-1})/2)$ 이고 탐색 영역은 원점을 중심으로 $[-1, +1]$ 화소의 범위이다. 본 단계에서 최소의 SAD를 가지는 움직임 벡터를 최종적으로 결정하게 된다. 본 논문에서 제안한 매크로 블록과 서브 매크로 블록 파티션 구성은 (그림 3)과 같다.



(그림 3) 변형계층 알고리즘 용 블록 파티션 구조

3.2 변형 알고리즘 성능 분석

기존의 계층적 방법은 조작된 값, 즉 휘도의 평균값으로 비교 블록을 생성한다. 이는 원본 이미지의 휘도 특성을 왜곡 시킬 수 있다. 또한 휘도의 평균값이 동일한 영상 일지라도 각 휘도 간 편차는 다를 수 있기에(단일 영상에서 명암의 차이가 현저한 경우) 이를 고려하지 않고 있다. 변형 방법은 해상도를 낮춘 부표본 블록을 생성 하면서도 원본 영상의 데이터를 손상 시키지 않는 방법을 제시한다.

기존 계층적 방법에서 부 표본(subsampling)을 만드는 것은 이미지의 크기를 줄이는 또 다른 접근법이다. 부 표본을 만드는 동안에 여러 개의 화소들로 이루어진 블록은 그 화소들 중의 하나로 대체된다. 이

방법은 휘도의 평균값을 이용하는 방법보다 화질은 떨어지지만 처리 속도는 더 빠르다. 변형 계층적 방법은 화질이 떨어지지 않으면서 속도를 빠르게 할 수 있는 부표본 방법이다.

또한 기존 계층적 방법은 높은 단계에서의 오류가 거의 항상 낮은 단계로 전파된다는 것이다. 이것은 결코 드문 현상이 아니다. 즉 많은 후보 블록들이 유사한 블록거리 측정값을 가질 것이고 실제적으로 정합된 블록은 다른 블록들에 비해 상위단계에서 좀 더 적합할 것이다. 그러나 0단계에서 반드시 더 적합한 것은 아니다. 반면 변형방법은 상하단계에서 동시에 탐색이 진행되므로 상위단계가 최종단계에 주는 악영향을 어느 정도 차단할 수 있고 단계의 깊이가 실질적으로는 2단계에 해당하므로 작은 물체의 경우에도 저해상도 영상에 반영할 수 있다. 즉 4×4 블록을 원본 영상으로 하여 탐색이 가능하다. 또한 16×8 , 8×16 , 8×4 , 4×8 블록을 이용한 직사각형 매크로 블록 단위로도 계층적 움직임 추정이 가능하다. 다만 기존 계층적 탐색 법의 단점인 다른 해상도의 영상을 저장하기 위한 메모리가 요구되는 점을 배가 시킬 수 있다.

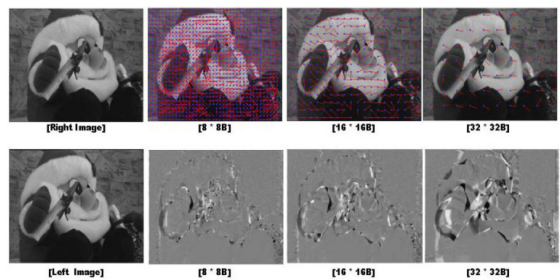
3.3 성능평가 실험 및 응용

성능을 평가하기 위해 실험에 사용한 영상은 비교적 질감의 정도가 중간 정도이며 밝기의 변화가 없는 폐쇄 영역을 부분적으로 많이 포함하고 있는 영상이다. (그림 4)의 영상의 크기는 320×240 픽셀이며, 각각 좌우 스테레오 영상과 계층된 시차지도를 나타내고 있다. 실험은 2.0GHz Pentium4, 1024Mbyte RAM 시스템에서 하였다.



(그림 4) 실험용 스테레오 영상과 기준 시차지도

(그림 5)는 변형 계층적 움직임 추정에 사용된 블록 사이즈에 따른 오차 프레임들을 보여주고 있다.



(그림 5) 블록 크기에 따른 오차 프레임

움직임 추정의 블록 사이즈에 따른 알고리즘 수행 시간은 <표 1>에 나타났다.

<표 1> 탐색 방법에 따른 움직임 벡터 수행시간

탐색 방법	전역 탐색	3단계 탐색	계층 (8x8)	계층 (16x16)	변형 (16x16)	계층 (32x32)
수행 시간	1.5000 sec	0.5000 sec	0.2350 sec	0.1250 sec	0.1220 sec	0.1090 sec

변형 계층적 탐색 법은 부표본의 일종으로 속도에서 약간의 우월성을 보였다. 휘도의 평균을 이용하는 계층적 방법보다는 많은 시간 단축을, 부 표본을 이용하는 계층적 방법보다는 정확도가 높아졌다. 그러나 변형 방법의 주된 목적은 속도를 유지하면서도 화질을 높이고자 하는 알고리즘이기에 앞으로 다양한 영상에 대한 추정 속도분석이 필요하고, 최악의 상황 등 여러 여건에서 화질의 측정이 요구된다.

4. 결론

계층적 움직임 추정 알고리즘이 처리시간을 단축시킬 수 있는 장점에 비하여 작은 물체를 포함하고 있는 영역에 대하여 부정확성을 보였고 저 해상도에서 얻는 이점 못지않게 모호한 결과를 유도하여 최종 단계에 전파되는 모순을 도출 하였다. 그래서 본 논문에서는 기존 계층 방법의 한 접근 방법인 부 표본 방법을 변형 제안하여 부표본 방법이 갖는 속도의 이점을 살리면서도 화질이 떨어지지 않도록 하였다. 기존 계층 방법이 휘도의 평균을 이용한 두 단계의 상위 단계로 구성되어지는 반면 변형된 방법은 상위 휘도와 하위 휘도로 구성된 같은 수준의 두 단계를 통해 적은 해상도의 비교 블록을 통한 움직임을 추정하였다. 이는 휘도의 평균값이라는 조작된 데이터에 의한 비교가 아닌 원본에 속한 데이터를 비교함으로써 정확도를 높이고 소형, 직사각형블록에서도 추정이 가능함을 보였다. 결과적으로, 기존 계층적 방법보다는 많은 시간 단축을, 부 표본을 이용하는 계층적 방법 보다는 정확도가 높아졌다. 다만 기존 방법이 갖는 부 표본 저장 공간 문제를 배가 시키는 문제점도 도출되었다. 향후 휘도의 편차가 큰 경우와 같은 다양한 악조건 영상에 대한 변형 알고리즘의 실험, 검증을 시행할 예정이며, 도출된 저장 공간 문제를 개선할 수 있는 방법과 휘도의 성분을 더욱 특성화할 수 있는 방안에 대하여 연구를 계속할 생각이다. 또한 처리속도를 더욱 단축시키는 움직임 추정 모듈의 최적화 방법에 대한 연구를 계속 진행할 계획이다. 본 논문에서

제시한 변형계층 움직임벡터 추정 방법은 유비쿼터스 환경에서 다양하게 적용이 가능할 것으로 생각된다.

참고문헌

- [1] <http://www.vqeg.org/> (Video Quality Experts Group).
- [2] Iain E.G. Richardson, H.264 and MPEG-4 [조중휘 · 손요한, 차세대 영상압축기술(홍릉과학출판사, 2004)].
- [3] D. Geiger, B. Ladendorf, and A. Yuille, "Occlusions and binocular stereo", Intl. Journal of Computer Vision, Vol.14, No.3, pp. 211-226, 1995.
- [4] ISO/IEC 14496 - 10 and ITU - T Rec. H.264, Advanced Video Coding, 2003.
- [5] F.Pereira and T.Ebrahimi(eds), *The MPEG -4 Book*, IMSC Press, 2002.
- [6] ISO/IEC 14495-1:2000 information technology - lossless and near-loss compression of continuous-tone still image : Baseline, (JPEG-LS).
- [7] ITU-T Recommendation, Information technology - coded representation of picture and audio information - progressive bi-level image compression, T.82(JBIG).
- [8] D. W. Kim, J. S. Choi and J. T. Kim, "Adaptive motion estimation based on spatiotemporal correlation," *Signal Processing: Image Commun.*, vol. 13, pp. 161-170, 1998.
- [9] MPEG-4 Industry Forum, <http://www.m4if.org>
- [10] ITU-T Recommendation H.263, Video coding for low bit rate communication, Version2, 1998.
- [11] J. C. Tsai, C. H. Hseigh, S. K. Weng, and M. F. Lai, "Block-matching motion estimation using correlation search algorithm," *Signal Processing: Image Commun.*, vol. 13, pp.119-133, 1998.
- [12] L. K. Liu and E. Feig, "A block-based gradient descent search algorithm for block motion estimation in video coding", *IEEE Trans. on Circuits Syst. Video Technol.*, Vol.6, pp. 419-423, 1996.
- [13] M. Brunig, W. Niehsen, "Fast full-search block matching," *IEEE Trans on Circuits and Systems for video Technology*, Vol.11, No.2, pp.241-247, Feb. 2001.