

스테레오 코딩을 위한 스테레오 영상의 밸런싱 방법

김종수*, 최종호**, 김태용*, 최중수*

A Balancing Method of Stereo Pairs for Stereo Coding

Jong-Su Kim*, Jong-Ho Choi**, Tae-Yong Kim*, Jong-Soo Choi*

요 약

3D 디스플레이 기술이 발전함에 따라 스테레오 영상의 전송시 요구되는 비트레이트의 감소가 절실히 필요하다. 하지만, 스테레오 영상은 서로 다른 카메라에 의해 취득되기 때문에 잠재적으로 서로 차이가 있고, 이것은 디스패리티 추정시 큰 오차를 유발할 수 있으며 전송될 비트레이트에 영향을 줄 수 있다. 따라서 스테레오 영상들 사이의 밸런싱이 필요하다. 스테레오 영상의 밸런싱을 위해, 본 논문에서는 히스토그램 Specification 방법과 타깃 영상의 국부정보, 스테레오 영상간의 오차 분포를 이용한다. 히스토그램 Specification 방법은 그레이레벨의 맵핑관계를 정의한다. 따라서 이를 통해 맵핑될 레벨의 맵핑 구간을 구할 수 있다. 그 구간에서, 맵핑될 기준영상의 히스토그램 분포와 스테레오 오차값의 분포는 서로 모양이 유사할 것이다. 그러나, 폐색된 영역이나 노이즈에 의해 그 모양이 변하므로 우리는 맵핑될 픽셀들을, 오차영상에서 그 픽셀들의 근방에서 구한 평균들과 오른쪽 영상(타깃 영상)에서 맵핑될 픽셀의 근방에서 구한 평균이 최소값을 갖는 위치값으로 맵핑한다. 제안된 방법은 실험에서 기존 방법보다 향상된 결과를 나타내는 것을 보여 준다.

▶ Keyword : 스테레오, 히스토그램 Specification, 밸런싱, 비트레이트

1. 서 론

최근 3D 디스플레이 장치들이 개발되고 보급화 되고 있다. 디스플레이 장치는 흑백에서 컬러, HDTV로 빠르게 발전되어 왔고 향후 3D 디스플레이 장치가 보급화 될 것으로 많은 전문가들이 예상하고 있다. 이에 따라, 영상 장치들 사이에 스테레오 영상 전송의 양은 증가할 것이다.

특히, 유비쿼터스 환경하에서는 이기종간의 영상 데이터 전송의 양은 폭발적으로 증가할 것이다. 하지만, 기존 표준 압축 방식하에서, 스테레오 영상은 모노 영상 전송에 비해 2배의 대역폭을 필요로 한다. 따라서, 한정된 대역폭을 효율적으로 사용하기 위해서는 효과적인 스테레오 영상 압축이 필요하다. 모노 영상 압축 시스템에서 DCT 보정된 영상과 입력된 영상간의 오차 영상은 전송 데이터의 양을 결정하는 중요한 요소이다.

* 제1저자 : 김종수

* 중앙대학교 첨단영상대학원 영상공학과, **강남대학교 전자공학전공

즉, 시간상의 중첩성을 이용하여 전송데이터의 양을 줄일 수 있다. 이와 같은 방식으로 스테레오 영상에서는 공간상의 중첩성을 이용할 수 있다. 즉, 두 영상 사이의 오차 영상을 전송함으로써 전송데이터의 양을 줄일 수 있다. 스테레오 영상의 중첩성은 스테레오 정합을 통해 계산할 수 있다. 스테레오 정합 방법은 정합 요소에 따라 특징기반법과 영역기반법으로 나눌 수 있다. 전체 영상의 변위값이나 회전, 크기값을 구하기 위해서는 특징기반법 보다는 영역기반법이 효과적이므로 본 논문에서는 밝기 정보와 잡음에 강건한 FFT 기반의 스테레오 정합 방법을 이용한다.

스테레오 영상은 서로 다른 장치에 의해 취득되기 때문에, 동일한 픽셀이라도 잠재적으로 오차가 있을 수 있다. 이것은 각 카메라의 노출정도나 포커스의 불일치, 카메라 설정등에 의해 유발될 수 있다. 따라서, 전송데이터의 양을 줄이기 위해서, 스테레오 영상의 밸런싱은 매우 중요하다. 이를 위해, 본 논문에서는 히스토그램 Specification 방법을 기반으로 밸런싱을 한다. 히스토그램 Specification 방법을 이용하면 타깃영상의 그레이레벨들의 맵핑 구간을 계산할 수 있다. 그 구간에서 맵핑될 기준 영상의 히스토그램 분포와 타깃 영상내에 맵핑될 픽셀들의 오차값의 분포는 그 모양이 서로 유사하다. 그러나 오차값의 분포는 폐색된 영역이나 노이즈의 영향을 받으므로 그 모양이 변하게 된다. 따라서 본 논문에서는 그 구간과 관련된 스테레오 오차값들의 분포 구간내에 존재하는 픽셀들의 근방에서 평균값들을 구하고 입력된 타깃 영상의 픽셀 근방에서 구한 평균값이 최소 오차를 갖는 위치값으로 입력 픽셀을 맵핑한다. 이를 통해 스테레오 영상의 밸런싱을 하게 된다.

본 논문의 구성은 다음과 같다. 2장에서 FFT 방법을 이용한 스테레오 정합과 기존 밸런싱 방법들에 대해 살펴본다. 3장에서는 제안한 밸런싱 방법에 대해 설명하고, 그 실험적 결과들을 4장에서 보여준다. 마지막으로 5장에서 결론을 맺는다.

2. 스테레오 영상 정합과 밸런싱

2.1 스테레오 영상 정합

스테레오 영상의 정합을 위해 다양한 많은 방법들이 제안되고 사용되고 있다. 가장 많이 이용되는 방법들은 크게 영상의 에지나 코너점을 이용한 특징점 기반의 방법, 영상 사이의 상관성을 이용한 방법, 그리고 주파수 공간상에서 처리되는 푸리에변환을 이용한 방법으로 나눌 수 있다. [1][2][3] 특히, 푸리에변환을 이용한 방법은 빠른 처

리와 노이즈에 강건한 특징이 있기 때문에 본 논문에서는 이 방법을 이용하여 스테레오 영상을 정합한다.

푸리에변환의 천이 성질은 스테레오 정합을 위해 이용될 수 있다. 한 쌍의 스테레오 이미지들을 각각 $f_1(x, y)$, $f_2(x, y)$ 라 하고, 그에 대응되는 푸리에변환을 $F_1(u, v)$, $F_2(u, v)$ 라 하자. 만약, 두 영상이 변위 성분 (x_0, y_0) 만 다르다고 하면, 두 영상 사이의 관계는

$$f_2(x, y) = f_1(x - x_0, y - y_0)$$

와 같고, 그에 대응된 푸리에변환 관계는

$$F_2(u, v) = \exp(-j2\pi(ux_0 + vy_0)) \times F_1(u, v)$$

이 된다. 천이 성질에 의해, 두 영상의 Cross-Power 스펙트럼은

$$G(u, v) = \frac{F_1(u, v) \times F_2^*(u, v)}{|F_1(u, v) \times F_2^*(u, v)|} = \exp(j2\pi(ux_0 + vy_0))$$

로 정의된다. $G(u, v)$ 의 역푸리에변환은 델타함수를 생성한다. 따라서, 두 영상 사이의 변위 성분 (x_0, y_0) 는 델타함수의 위치로부터 쉽게 추정할 수 있다. 두 영상 사이의 상대적인 회전 및 크기 성분도 이와 유사하게 계산할 수 있다.

두 영상 사이에 회전 성분 θ_0 와 크기 성분 a 가 존재한다고 하면 두 영상 사이에는 다음과 같은

$$f_2(x, y) = \frac{1}{a} f_1(x \cos \theta_0 + y \sin \theta_0, -x \sin \theta_0 + y \cos \theta_0)$$

관계가 성립한다. 위 수식에서 푸리에변환의 그 크기를 polar 변환을 하게 되면,

$$M_2(\rho, \theta) = M_1(\log \rho - \log a, \theta - \theta_0)$$

와 같이 나타낼 수 있다. 여기서 M_1 , M_2 는 F_1 과 F_2 의 크기들을 polar 변환한 것을 나타낸다. 회전성분과 크기 성분은, 변위 성분에서와 같이, M_1 과 M_2 의 Cross-Power 스펙트럼을 구하게 되면 간편히 계산된다.

스테레오 영상은 회전과 크기성분을 먼저 구한 후, 계산된 회전 및 크기성분으로 영상을 변환하여 변위성분을 구하고, 계산된 변위성분으로 재차 변환하면 정합된 영상을 구하게 된다.

2.2 기존 밸런싱 방법

스테레오 영상의 밸런싱은 선형 또는 비선형 방법을 이용할 수 있다. 전통적인 선형 밸런싱 방법은 평균과 표준편차를 이용한다.[4][5] 이 방법은 빠른 처리의 장점을 가진 반면, 평균 추정 에러와 유니모달인 경우만 고려하였기 때문에, 큰 오차를 유발할 수 있다. 비선형 방법은 국부적 정보를 이용한 방법과 전체 영상 정보를 이용한 방법으로 분류할 수 있다. 국부정보를 이용한 방법은 영상 개선을 위해 많이 사용되지만 계산적 복잡도 때문에 스테레오 영상 밸런싱에는 적합하지 않다. 전체 영상 정보를 이용한 방법으로는 대표적인 히스토그램 Specification 방법이 있다. 본 논문에서는 국부적 정보와 전체 영상 정보를 같이 이용한다.

히스토그램 Specification은 히스토그램 맵핑 관계에 따라 많은 방법이 제안되었다. SML(Single Mapping Law)과 GML(Group mapping Law)은 초기에 제안된 다이렉트 히스토그램 Specification 방법이다.[6] 이 방법들은 일대일(one-to-one) 맵핑(one-to-one)과 다대일(many-to-one) 맵핑을 하기 때문에 변환속도는 빠르지만 매칭에러가 확산되고 누적되기 때문에 큰 에러를 야기할 수 있다. 다이렉트 방법들의 매칭에러를 줄이기 위해, 다이내믹 프로그래밍을 통하여 다중(one-to-many) 맵핑도 고려한 DHW (Dinamic Histogram Warping) 방법이 제안되었다.[7] 이 방법을 이용하면 매칭에러는 줄일 수 있지만, 디지털 영상에서 정확한 히스토그램 Specification은 수행하지는 못한다. Dinu Coltuc 등은 [8]에서 strict Ordeing을 통해 디지털 영상의 정확한 히스토그램 명확화를 수행하였다. Strict Ordering 방법은, 히스토그램에서 다중(One-to-many) 맵핑인 경우 분포화를 정확히 하기 위해, 같은 화소를 갖는 픽셀의 주변 영역을 확장해가며 평균값을 계산하여 Ordering 하는 방법이다. 이 방법은 변환할 영상에서 중첩영역의 모든 화소들을 Ordering하기 때문에 그 처리 속도가 느리고 중첩되지 않은 영역의 밸런싱을 할 수 없는 단점을 가지고 있다.

3. 스테레오 영상의 밸런싱

본 논문에서는 스테레오 영상의 밸런싱을 위해 중첩된 영역의 히스토그램을 이용한다. 만약, 스테레오 영상 정합이 정확하게 이루어졌다면, 중첩된 영역에서 왼쪽영상(기준영상)과 오른쪽영상(타깃영상)의 픽셀들은 서로 대응된다. 스테레오 영상들 사이에 폐색된(occluded region) 영역이 거의 없고 밸런스 되어 있다고 하면, 중첩된 영역에

서 기준영상과 타깃영상의 히스토그램은 $hr \cong h_l$ 와 같이 서로 유사할 것이다. 그러나 스테레오 영상은 서로 다른 장치에 의해 취득되기 때문에, 히스토그램 분포는 변하게 된다. 히스토그램 specification 방법은 중첩된 영역에서 기준영상의 히스토그램과 같은 분포를 갖게 타깃영상의 그레이레벨을 변환하는 방법이다. 타깃영상의 그레이레벨 a 가 기준영상의 그레이레벨 m 으로 맵핑된다고 하면, 밸런싱 함수는 $B(a)=m$ 이 된다. 따라서 정확한 $B(a)$ 를 추정하는 것이 중요하다.

타깃영상의 한 그레이레벨은 기준영상의 하나 또는 여러 개의 레벨들로 맵핑될 것이다. 그 맵핑 구간을 나타내는 함수를 $w(i)$, $i=0, 1, \dots, 255$ 라 하자. 따라서 타깃영상의 한 레벨 a 가 n 개의 레벨들로 맵핑된다고 하면, $w(a)=n$ 이 된다. 그러므로, 타깃영상의 모든 레벨들은 항상 $w(i) \geq 1$ 이고, 모든 픽셀들은 일대일(one-to-one: $w(i)=1$) 맵핑이나 다중맵핑(one-to-many: $w(i) > 1$) 관계에 있다고 할 수 있다. $w(i)=n$ 인 경우, 타깃영상의 i 값을 갖는 픽셀들은 $\{k(i), k(i)+1, \dots, k(i)+n-1\}$ 중에 한 값으로 맵핑될 것이다. 여기서 $k(i)$ 는 맵핑될 가장 작은 값을 나타내고

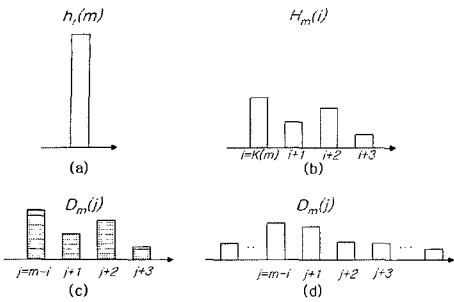
$$k(i) = \begin{cases} 0, & \text{if } i = 0 \\ j+1, & \text{if } (c_r(i-1) - c_l(j)) = 0 \text{ and } 0 < (c_r(i-1) - c_l(j-1)) \\ j, & \text{if } (c_r(i-1) - c_l(j)) < 0 \text{ and } 0 < (c_r(i-1) - c_l(j-1)) \end{cases}$$

와 같이 구할 수 있다. 여기서 c_r 과 c_l 은 각각 h_r 과 h_l 의 누적 히스토그램 분포를 나타낸다. $k(i)$ 에 의해 $w(i)$ 는

$$w(i) = \begin{cases} k(i+1) - k(i), & \text{if } c_l(k(i+1)-1) = c_r(i) \\ k(i+1) - k(i) + 1, & \text{else} \\ 255 - k(i) + 1, & \text{if } i = 255 \end{cases}$$

이 된다. 여기서 $w(i) > 1$ 인 경우, 타깃영상의 그 픽셀들은 다중값들로 분리해야 한다. 이를 위해, 본 논문에서는 기준영상과 타깃영상간의 오차값의 분포를 이용한다. $H_m(i)$, $k(m) \leq i \leq k(m) + w(m) - 1$ 가 $hr(m)$ 이 분리될 기준영상의 히스토그램 분포라 하고, $D_m(n)$, $n = m-i$ 가 타깃영상에서 그레이값이 m 인 픽셀들과 그에 대응되는 기준영상 픽셀들간의 오차값의 분포라 하자. 그러면, $\sum H_m(i) = h_r(m)$ 이 되고 $i=n$ 이 된다. 만약, 노이즈와 폐색된 영역이 없다 하면, 우리는 다음과 같은 수식을 얻을 수 있다.

$$\sum_n D_m(n) = \sum_i H_m(i) = h_r(m)$$



〈그림 1〉 히스토그램 다중 맵핑

(a) 맵핑될 타깃영상의 한 레벨의 히스토그램, (b) 분리될 기준의 영상의 히스토그램 분포, (c) 노이즈나 폐색된 영역이 없는 경우의 오차값 분포, (d) 노이즈나 폐색된 영역이 있는 경우의 오차값 분포.

따라서 오차 분포를 $hr(m)$ 을 분리하기 위해 사용할 수 있다. 그러나, 노이즈나 폐색된 영역에 의해 오차 분포의 모양은 변하므로, 우리는 이것을 고려하기 위해, $n < m-i$ 이거나 $m-i+w(m) \leq n$ 인 픽셀들은 노이즈나 폐색된 영역의 픽셀로 간주하고, 오차영상에서 $m-i \leq n < m-i+w(m)$ 인 픽셀들의 이웃하는 근방의 평균값들을 구한다.

$$M_m(n) = \frac{1}{D_m(n) \times 9} \sum_{x=y=0}^{x=N_x-M} \sum_{y=0}^{y=N_y-1} (I_r(x+i, y+j)), \text{ if } d_m(x, y) = n - (w(m) - 1)$$

그 평균값들 $M_m(n)$ 는 $I_r(x, y) = m$ 인 픽셀들을

$$I_r(x, y) = \arg \min_n |M - M_m(n)|$$

와 같이 분리할 수 있다. 여기서

$$M = \frac{1}{9} \sum_{i=j=1}^{i=j=1} I_r(x+i, y+j)$$

이다. 요약하자면, 타깃영상은

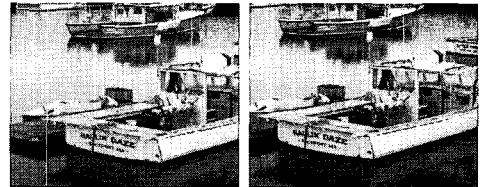
$$B(I_r(x, y)) = \begin{cases} k(m), & \text{if } w(m) = 1 \\ \arg \min_n |M - M_m(n)|, & \text{if } w(m) \neq 1 \end{cases}$$

으로 밸런싱 된다.

4. 실험 및 결과

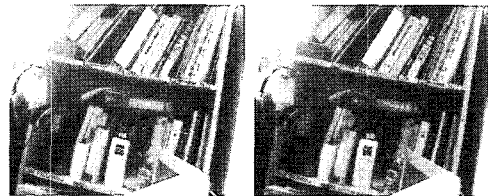
제안한 방법의 검증을 위해, 본 논문에서는, 서로 다른 특성을 갖는 카메라에서 취득된 영상을 가지고 실험을 하였다. 배 영상은 카메라특성을 달리하여 취득한 영상이고,

실험실 영상은 웹 카메라를 이용하여 취득 영상이다. 이때, 카메라들의 세팅 파라미터인 밝기, 포화, 콘트라스트, 감마 및 화이트 밸런스 값을 다르게 하여 취득하였다. 그림 2는 취득된 영상들을 보여 주고 있다.



(a) 배 기준 영상

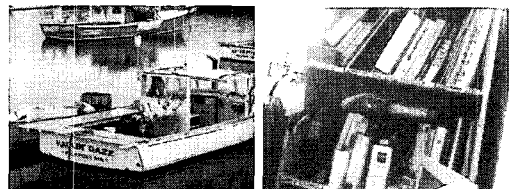
(b) 배 타깃 영상



(a) 실험실 기준 영상

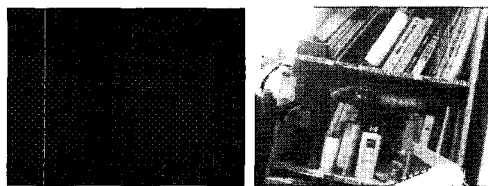
(b) 실험실 타깃 영상

〈그림 2〉 실험 영상들

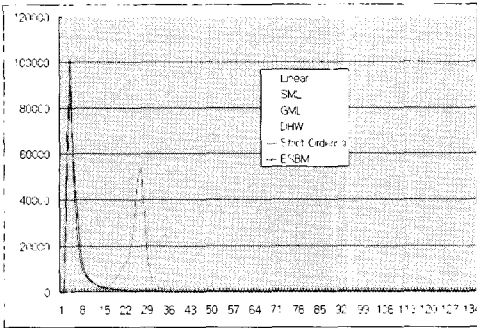


〈그림 3〉 밸런싱된 타깃 영상들

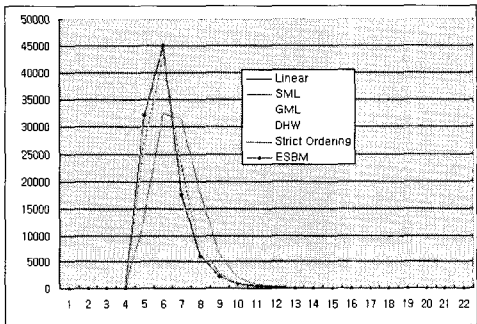
본 논문에서 제안한 방법을 이용하여 밸런싱된 영상은 그림 3과 같고, 시각적으로 기준 영상과 차이를 분간하기 어려운 밸런싱된 결과를 볼 수 있다. 그리고 그림 4는 밸런싱된 영상을 이용하여 디스패리티 추정 결과와 디스패리티 보상된 결과를 보여주고 있다. 제안한 방법과 기존 방법들을 비교하기 위해 본 논문에서는, 중첩된 영역에서, 밸런싱된 영상과 기준 영상간의 오차영상을 구하여 그 오차 분포를 측정하였다. 그림 5는 각 방법들로 측정된 오차 분포를 보여주고 있다. 기존 방법들에 비해 제안한 방법의 오차 분포가 작은 구간에서 분포함을 볼 수 있다.



〈그림 4〉 디스패리티 추정 결과와 보상된 영상



(a)배 영상의 오차분포



(b)실험실 영상의 오차분포
(그림 5) 측정된 오차 분포의 비교

또한, 본 논문에서는 PSNR을 계산하여 기존 방법들과 비교하였다. 표 1에서는 PSNR 결과들을 보여주고 있으며, 제안된 방법과 기존방법들간에 PSNR의 차이가 크지 않은 것을 알 수 있다.

〈표 1〉 PSNR의 비교

배 영상	배영상	실험실영상
선형 방법	38.5780	32.8983
SML 방법	32.7486	28.6538
GML 방법	40.5373	33.0115
DHW 방법	38.4131	30.9755
Ordering 방법	40.4712	33.0012
제한한 방법	40.5385	32.9703

5. 결론

본 논문에서는 스테레오 영상의 뱀런싱 방법을 제안하였다. 제안된 방법을 이용하면, 실험 결과에서 살펴본 바와 같이, 기존 방법에 비해 향상된 결과를 얻을 수 있다. 제안한 방법에 의해 뱀런싱된 영상과 기준 영상간의 오차 영상은 기존 방법들에 비해 오차 분포가 매우 작은 구간에

집중됨을 실험을 통해 살펴보았다. 이는, 전송될 비트레이트의 크기에 영향을 주는 것은 자명한 일이다. 따라서 본 논문에서 제안된 방법은 향후 스테레오 코딩 시스템에 적용되어, 전송 비트레이트의 감소를 유도 할 수 있다.

참고문헌

- [1] E. De Castro and C. Morandi, "Registration of translated and rotated images using finite Fourier transforms," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, pp. 700-703, May 1987
- [2] B.S. Reddy and B.N. Chatterji, "An fft-based technique for translation, rotation, and scale-invariant image registration", *IEEE Transactions on Image Processing*, vol.5, no.8, pp.1266-1271, August 1996.
- [3] R. Eustice, O. Pizarro, H. Singh, and J. Howland, "UWIT: underwater image toolbox for optical image processing and mosaicking in Matlab," in *Proc. Intl. Symp. Underwater Tech*, Tokyo, Japan, pp.141 - 145, April, 2002,
- [4] A. Mancini, "Disparity estimation and intermediate view reconstruction for noble applications in stereoscopic video", *Master Thesis, McGill University*, February, 1998.
- [5] I. Patras, N. Alvertos, G. Tziritas, "Joint disparity and motion field estimation in stereoscopic image sequences", *Proc. IEEE ICPR*, pp.359 - 363, 1996
- [6] Y.J.Zhang, "Improving the Accuracy of Direct Histogram Specification", *Electronics Letters*, Vol.28 No3, January 1992.
- [7] I.J. Cox, S. Roy, S.L. Hingorani, "Dynamic Histogram Warping of Image Pairs for Constant Image Brightness", *International Conference on Image Processing(ICIP)*, Vol. 2, 1995.
- [8] Dinu Coltuc, Philippe Bolon, and Jean-Marc Chassery, "Exact Histogram Specification", *IEEE Transactions on Image Processing*, Vol.15 No.5, May 2006.