

## 다중 카메라의 동적인 선택 알고리즘 개발

장석우<sup>○</sup>, 최현준<sup>\*</sup>, 이숙윤<sup>\*\*</sup>

<sup>○</sup>안양대학교 디지털미디어학과

<sup>\*</sup>목포해양대학교 전자공학과

<sup>\*\*</sup>메트로비전

e-mail:swjang@anyang.ac.kr<sup>○</sup>, hjchoi@mmu.ac.kr<sup>\*</sup>, uni7@gmail.com<sup>\*\*</sup>

## Developing a Dynamic Selection Algorithm in Multiple Cameras

Seok-Woo Jang<sup>○</sup>, Hyun-Jun Choi<sup>\*</sup>, Suk-Yun Lee<sup>\*\*</sup>

<sup>○</sup>Dept. of Digital Media, Anyang University

<sup>\*</sup>Department of Electronic Engineering, Mokpo National Maritime University

<sup>\*\*</sup>MetroVision

### ● 요약 ●

본 논문에서는 카메라가 여러 개 존재하는 다중의 카메라 환경에서 주변의 환경에 최적으로 적합한 카메라를 동적으로 선택하는 알고리즘을 제안한다. 제안된 알고리즘에서는 초기의 입력영상을 받아들인 후, 이 영상으로부터 주위의 환경을 가장 잘 표현할 수 있는 특징인 밝기와 텍스처 특징을 추출한다. 그리고 이전 단계에서 추출된 밝기와 텍스처 특징값들을 가장 잘 반영할 수 있는 카메라를 선택하는 규칙을 생성함으로써 주위 환경에 맞는 카메라를 자동으로 선택해 준다. 본 논문의 실험결과에서는 제안된 방법이 여러 가지 환경에서 잘 동작하며, 결과적으로 주위 환경에 적합한 카메라의 선택을 통해 보다 정확한 3차원의 정보를 추출함을 보여준다.

키워드: 선택 알고리즘(selection algorithm), 다중 카메라(multiple camera), 특징(feature)

### I. 서론

3차원 디스플레이 기술은 2차원의 영상에 깊이 정보를 추가하고 있어 생동감 및 현실감을 느낄 수 있게 하는 차세대 신기술이다. 현재 3차원 디스플레이 산업에 대한 시장의 관심은 가히 폭발적이라 할 수 있으며, 표준화 및 기술의 안정성 확보 등 차후 해결해야 할 부분도 많이 있으나 이런 문제에도 불구하고 3D가 차세대 주력 산업이라는 데에 이견은 없는 실정이다[1,2].

이런 3차원적인 입체를 효과적으로 구현하기 위한 핵심 요소기술 중의 하나는 3차원 카메라 처리 기술이다. 기존의 대부분의 3차원 카메라는 한 종류의 카메라를 고정적으로 사용하여 장면을 촬영하므로 영상의 품질에는 한계가 있을 수 밖에 없다. 따라서 본 논문에서는 수동형 방식의 카메라[3]와 능동형 방식의 카메라[4]를 결합하여 다중의 카메라를 구성하고, 주변 환경의 복잡도에 따라서 다중 카메라의 수동형과 능동형 카메라를 적응적으로 선택함으로써 3차원의 영상정보를 보다 정확하게 추출할 수 있는 알고리즘을 제안한다.

제안된 방법에서는 먼저 다중의 카메라를 구성한 후 초기 영상을 받아들이고, 이 입력영상으로부터 주변의 환경을 대표할 수 있는 주요 특징인 밝기 특징과 텍스처 특징을 추출한다. 그런 다음, 추출된 밝기와 텍스처 특징값을 가장 잘 수용할 수 있는 카메라를 선택하는 규칙을 생성함으로써 다중의 입체 카메라를 동적으로 선택하고, 이를 통해 품질이 보다 우수한 영상을 획득한다.

### II. 특징추출

본 논문에서는 주위 환경의 상태를 대표하는 특징으로서 밝기와 텍스처 특징을 정의하여 사용한다. 그리고 잡음에 의한 영향을 줄이기 위해서 입력영상을  $N \times N$  화소 크기의 블록으로 분할한 후, 블록 단위로 이런 특징들을 추출한다.

밝기 특징으로는 입력영상의 RGB 칼라공간을 색상(Hue), 명도(Intensity), 채도(Saturation)로 용이하게 변환할 수 있는 HIS 칼라공간[5]으로 변환한 후 I 값을 활용한다. 그림 1은 HIS 칼라공간을 기하학적으로 보여준다.

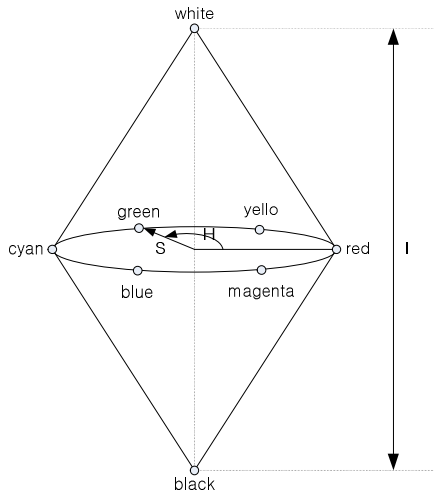


그림 1. HSI 칼라 모델  
Fig. 1. HSI color model

본 논문에서 t시점의 영상에 대한 밝기 특징은 식 (1)과 같이 정의된다. 식 (1)에서 N는 블록의 가로 또는 세로의 길이를 나타내고, P는 영상에서 샘플링된 블록의 전체 개수를 의미하며,  $R_i$ 는 i번째 블록 내의 위치 좌표들의 집합을 나타낸다. 그리고  $I_t(x,y)$ 는 영상 (x,y) 위치에서의 밝기 값을 의미한다. 식 (1)에서 t시점 영상의 밝기 특징은 0에서 1사이의 값을 가지도록 정규화(normalization)된다.

$$F_{bright}(t) = \frac{1}{256 \times N^2 \times P} \sum_{i=1}^P \sum_{(x,y) \in R_i} I_t(x,y) \quad (1)$$

텍스처 특징으로는 다른 텍스처 특징보다 우수한 텍스처 구분력을 가진 것으로 알려져 많이 사용되고 있는 가보 웨이블릿 변환[6]을 텍스처 특징으로 사용한다.

일반적으로, 가보 웨이블릿 변환은 가보 함수를 모함수(mother function)로 사용하는 웨이블릿 변환이다. 가보 함수는 복잡한 시누소이드(sinusoid) 함수에 의해 변조된 가우시안 함수로 2차원 가보 함수  $g(x,y)$ 는 식 (2)와 같이 정의된다.

$$g(x,y) = \left( \frac{1}{2\pi\sigma_x\sigma_y} \right) \exp \left( -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi j Wx \right) \quad (2)$$

식 (2)에서  $\sigma_x$ 와  $\sigma_y$ 는 각각 x와 y축으로의 표준편차이고, W는 관심의 대상이 되는 기본 주파수이다.

$$F_{gabor}(t) = \sum_{x,y \in R} \sum_{k=x-w}^{x+w} \sum_{l=y-w}^{y+w} \frac{I_t(k,l) \cdot g_{mn}(k-x,l-y)}{N^2} \quad (3)$$

식 (3)의 w는 (x,y)를 중심으로 하는 윈도우의 반경을 나타낸

다. 본 논문에서는 i번째 블록이 대표하는 영역의 모든 좌표들 (x,y)에 대해 이를 중심으로 하는 윈도우를 사용해 계산되는 가보 웨이블릿 계수들의 평균을 각 블록의 텍스처 특징으로 사용한다.

### III. 카메라의 선택 방법

주변 환경의 복잡도를 대표하는 밝기와 텍스처 특징을 정의한 후에는 이들 특징을 효과적으로 결합하는 통합척도(integrated measure)를 정의하고, 정의된 통합척도를 기반으로 동적인 카메라의 선택 알고리즘을 개발한다.

본 논문에서는 t 시점의 입력영상에 대해 주변 환경의 복잡도를 효과적으로 측정하는 통합척도  $\Phi(t;\alpha,\beta)$ 를 밝기 특징과 텍스처 특징의 가중화된 합을 이용하여 식 (4)와 같이 정의한다. 식 (4)에서 통합척도는 밝기와 텍스처 특징을 나타내는 두 항의 가중화된 합(weighted sum)을 이용해 정의된다.

$$\Phi(t;\alpha,\beta) = \alpha \times \left( 1 - \left| \frac{F_{bright}(t)}{\|F_{bright}(t)\|} - \frac{1}{2} \right| \right) + \beta \times \left( \frac{F_{gabor}(t)}{\|F_{gabor}(t)\|} \right) \quad (4)$$

where  $0 \leq \alpha, \beta \leq 1, \alpha + \beta = 1$

주변 환경의 복잡도를 측정하기 위한 통합척도를 정의한 후에는 다음과 같은 규칙을 이용하여 다중의 카메라를 동적으로 스위칭한다. 즉, 통합 척도가 임계치 TH 이상일 경우에는 다중 카메라의 수동형 카메라를 활성화하고, 그렇지 않을 경우에는 능동형 카메라를 활성화한다. 다시 말해, 주변 환경의 밝기가 어둡거나 밝을 경우, 그리고 텍스처가 빈약할 경우에는 능동형 카메라를 활성화하고, 주변 환경의 밝기와 텍스처가 빈약하지 않을 경우에는 수동형 카메라인 스테레오 카메라를 활성화한다.

$IF(\Phi(t;\alpha,\beta) \geq TH)$   
 THEN select passive camera  
 ELSEIF ( $\Phi(t;\alpha,\beta) < TH$ )  
 THEN select active camera  
 END

### IV. 실험결과

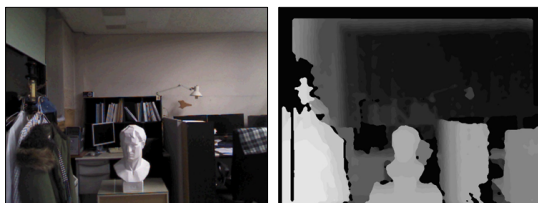
본 논문에서는 수동형 카메라인 3차원 스테레오 카메라로는 Pointgrey사의 Bumblebee2 카메라, 그리고 그림 2와 같은 싱글 카메라 2대를 획득 영상의 크기와 전송속도의 향상을 위해 IEEE1394단자에 평행 방식으로 배열하여 스테레오 영상을 획득하였다. 그리고 본 논문에서 사용한 능동형 카메라는 마이크로소프트사의 Kinect를 사용하였다.



그림 2. 수동형 카메라  
Fig. 2. Passive camera

그림 3과 그림 4는 조명과 텍스처가 다양하게 존재하는 입력영상을 받아들인 후 본 논문에서 제안된 동적인 카메라의 선택 알고리즘을 적용하여 추출한 변위 맵(disparity map)의 예를 보여준다. 그림 3은 조명이 어둡고 텍스처가 빈약한 환경에서 촬영한 영상에 대한 실험인데, 제안된 선택 알고리즘으로 인해 수동형 카메라 대신 능동형 카메라가 선택되어 변위 맵이 보다 정확성 있게 추출되었다. 그리고 그림 4는 조명이 어둡고 텍스처가 매우 빈약한 환경에서 촬영한 영상에 대한 실험인데, 부 정확하게 추출된 변위 맵의 예를 보여준다. 그림 3에서는 조명이나 텍스처 특징이 부족한 환경에서 능동형 카메라 대신 수동형 카메라가 선택되었을 경우에 3차원의 깊이 정보를 올바르게 측정할 수가 없으므로 정확도가 낮은 변위 맵이 획득되는 것을 시각적으로 확인할 수 있었다.

본 논문에서 제안된 다중 카메라의 동적인 선택 알고리즘을 사용한 방법을 전체적으로 판단해 볼 때, 시간적인 측면에서는 제안된 방법이 능동형 카메라 보다는 성능이 높았으며, 수동형 방식에 비해서는 약간 낮은 성능을 보였다. 그리고 정확도 측면에서는 제안된 알고리즘이 수동형 카메라를 사용한 방식보다는 우수하였고, 능동형 카메라를 이용한 방식에 비해서는 다소 낮은 정확도를 가짐을 시각적으로 확인하였다. 결과적으로, 제안된 동적인 카메라의



(a) Input image1 (b) Disparity map1

그림 3. 입력영상과 변위맵(1)

Fig. 3. Input image and disparity map(1)



(a) Input image2 (b) Disparity map2

그림 4. 입력영상과 변위맵(2)

Fig. 4. Input image and disparity map(2)

선택 알고리즘은 정확도와 시간적인 측면에서 기존의 수동형 또는 능동형 방식을 대체할 수 있는 좋은 성능을 가진 것으로 판단되었다.

## V. 결론

본 논문에서는 수동형 카메라와 능동형 카메라를 결합하여 다중의 입체 카메라를 구성하였다. 그런 다음, 주변 환경의 복잡도를 나타내는 밝기 특징과 텍스처 특징의 정도에 따라서 다중 카메라의 수동형과 능동형 카메라를 적응적으로 선택함으로써 3차원의 영상 정보를 보다 강건하게 추출할 수 있는 메커니즘을 제안하였다. 또한, 실험에서는 본 논문에서 제안한 동적인 선택 알고리즘이 여러 가지 조건에서 신뢰성이 있게 동작하였으며, 능동형 또는 수동형 카메라만을 단독적으로 사용하는 기존의 3차원 카메라를 대체할 수 있는 다중 카메라에 실제적으로 적용 가능성을 확인하였다.

향후에는 보다 다양한 종류의 입체 카메라를 추가하여 실험할 예정이며, 주변의 환경 변화에 보다 적응적으로 대응하면서 고품질의 입체 영상을 획득할 수 있는 연구를 수행할 예정이다.

## Acknowledgements

본 논문은 중소기업청에서 지원하는 2012년도 산학연공동기술개발 사업(No. C0027700)의 연구수행으로 인한 결과물임을 밝힙니다.

## 참고문헌

- [1] H.-J. Choi, Y.-H. Seo, S.-W. Jang, and D.-W. Kim, "Analysis of Digital Hologram Rendering Using a Computational Method," *Journal of Information and Communication Convergence Engineering*, Vol. 10, No. 2, pp. 205-209, Jun. 2012.
- [2] A. Smolic, "3D Video and Free Viewpoint Video— From Capture to Display," *Pattern Recognition*, Vol. 44, No. 9, pp. 1958-1968, Sep. 2011.
- [3] F. Dornaika and A. D. Sappa, "A Featureless and Stochastic Approach to On-Board Stereo Vision System Pose," *Image and Vision Computing*, Vol. 27, No. 9, pp. 1382-1393, Aug. 2009.
- [4] S.-Y. Lee and S.-W. Jang, "A Study on Image Processing for the Accuracy Improvement of 3D Recovery," In *Proc. of the Conference of the Korea Society of Computer and Information*, Vol. 20, No. 1, pp. 193-195, Jan. 2012.
- [5] K. Bunte, M. Biehl, M. F. Jonkman, and N. Petkov, "Learning Effective Color Features for Content Based Image Retrieval in Dermatology," *Pattern Recognition*, Vol. 44, No. 9, pp. 1892-1902, 2011.
- [6] S.-W. Jang, M.-A. Chung, and G.-Y. Kim, "Face Model-based Image Registration for Generating Facial Textures," In *Proc. of the ACM Research in Applied Computation Symposium*, pp. 121-125, Oct. 2012.