

영역분할과 움직임 정보를 이용한 깊이맵 생성 기법

*김수동 **안재우 ***서영호 ****김동욱 *****유지상
광운대학교

[*superksd@naver.com](mailto:superksd@naver.com) [**skai9@kw.ac.kr](mailto:skai9@kw.ac.kr) [***yhseo@kw.ac.kr](mailto:yhseo@kw.ac.kr) [****dwkim@kw.ac.kr](mailto:dwkim@kw.ac.kr)
[*****jsyoo@kw.ac.kr](mailto:jsyoo@kw.ac.kr)

Depth map generation method using segmentation and motion information

*Kim Sudong **Ahn Jaewoo ***Seo Young Ho ****Kim Dong Wook *****Yoo Jisang
Kwangwoon University

요약

본 논문에서는 영역 분할과 영상의 움직임 정보를 이용한 깊이맵 생성에 관한 기법을 제안하였다. 2D/3D 변환 알고리즘에서 2차원 영상에서 얻은 깊이 정보는 2차원 영상을 3차원 영상으로 변환 가능하게 하는 핵심 기술이 된다. 영역을 분할하고 계산되어진 움직임 값 (intensity)을 분할된 각 영역에 부여함으로써 깊이맵을 얻을 수 있다. 본 논문에서는 초기 단계에서 영역을 분할한 뒤, 입력 영상을 그룹화 하여 양방향 탐색을 통한 움직임 추정 연산을 수행토록 하여 보다 정확한 깊이 정보를 획득하고, 최종적으로 얻은 결과에 각 화소에 해당 되는 확률적 통계에 의한 후처리 기법을 사용하였다. 보다 정확한 깊이정보를 영역별로 지정하고, 후처리 기법을 사용함에 따라 보다 신뢰도 높은 깊이맵 영상을 생성할 수 있었다.

1. 서론

최근 제임스카메론 감독의 3D영화 “AVATAR”를 필두로 3D에 대한 관심이 높아지고 있고, 3DTV 및 3D방송, 3D영화에 관한 관심이 폭발적으로 증가하는 추세에 있다. 수요는 증가하지만 공급은 그에 비해 턱없이 부족한 실정이지만 현재 3D방송기술은 많은 부분에서 발전해야 할 필요가 있고, 그에 기인한 3D콘텐츠 부족에 따른 사용자들의 불만을 덜어주기 위한 일환으로 기존의 2D영상을 3D로 변환하는 방법이 해결책으로 꾸준히 제시되고 있다.

2D/3D 변환이란 2D 영상을 분석하여 정보를 추출해내고, 얻은 정보로부터 깊이맵을 생성한 뒤 3D 영상으로 변환하는 과정을 말한다. 2차원 평면의 영상으로부터 영상의 특징을 파악해내기란 쉬운 문제가 아니므로 깊이 정보가 존재하지 않는 2차원 영상으로부터 영상의 특징을 분석하여 깊이맵을 생성하는 것이 2D/3D 변환의 핵심 기술이라고 할 수 있다.

깊이맵을 생성하는데에는 여러 가지 방법이 존재한다. 객체의 외곽선 성분을 추출하여 수행하는 방법, 객체의 움직임 정보를 추출하여 그에 따른 깊이값을 부여하는 방법, 객체를 추출한 뒤 추적하여 깊이를 부여하는 방법 등 많은 방법들에 대한 연구가 진행중에 있다.

본 논문에서는 영역분할과 추정된 움직임 정보를 이용한 깊이맵 생성기법을 제안하였다. 영역분할은 영상을 어떠한 특징에 의하여 구분 짓고 각각 다른 영역으로 지정하는 것을 말하는데, 본 논문에서는 화소별 색상값의 특징들을 비교함으로써 영상을 비슷한 색상들의 분포를 가지는 각기 다른 여러 영역들로 분할하였다. 영역 분할은 객체의 외곽선을 뚜렷하게 분할하므로 차후 깊이맵을 부여하는 과정에서의 오류들을 줄여주는데 그 목적이 있다. 분할된 영역은 그룹화된 영상들간

의 움직임 값에 기인하여 각각 같은 깊이값 들을 갖는 서로 다른 여러 영역들로 구분되어 진다. 최종적으로 깊이맵을 생성하는 과정에서 전후 영상간의 차이들을 기준 영상 앞, 뒤영상과의 깊이맵의 평균값을 부여함으로써 보다 신뢰도 높은 깊이맵을 얻을 수 있도록 하였다.

2. 제안하는 알고리즘

가. 영역 분할 (segmentation)

본 논문에서는 영상의 색상 정보를 이용하여 영역을 분할하였다. 색상 정보는 동질의 영역에 대한 정보를 포함하고 있는 정보로써 색상의 유사도는 영역을 분할하기 위한 척도로 사용되어진다.

영역 분할은 초기화 단계, 영역 융합, 최소 화소 이하의 영역 처리 단계로 나누어진다. 초기화 단계에서는 2차원 영상의 화소 각각이 하나의 영역으로 초기화된다. 영역 융합 단계는 유사한 색상을 가진 영역을 하나의 영역으로 융합하는 과정이다. 유사한 색상의 구분을 위한 척도로 본 논문에서는 위도 값과 채도 값을 사용하였으며 유사도를 비용 함수로 사용한다. Y, U, V 각각의 비용 값은 식 (1)을 통해 구하며 이 값이 작을수록 두 영역이 비슷한 하나의 영역임을 의미한다.

$$\begin{aligned} C_Y(R_i, R_j) &= \{Y(R_i) - Y(R_j)\} \\ C_U(R_i, R_j) &= \{U(R_i) - U(R_j)\} \\ C_V(R_i, R_j) &= \{V(R_i) - V(R_j)\} \end{aligned} \quad (1)$$

R_i 는 영상에서 i 번째 영역을 의미한다. $Y(R_i)$ 는 영역 R_i 의 Y

성분의 평균값, $C_r(R_i, R_j)$ 는 영역 R_i 과 R_j 사이의 Y 성분의 비용값을 의미하고 U, V에 관한 식도 이와 같다.

식(1)에서 계산된 Y, U, V 성분의 각각의 계산된 비용값이 모두 설정된 허용 오차보다 작다면 영역을 융합하는 과정이 수행된다. 허용 오차값은 잘못된 융합의 발생을 막기위해 임의의 작은 값 (0 ~ 5)으로 설정 하였다.

영역 융합 과정을 거치고 나면 영역이 과도하게 분할된 결과를 얻을 수 있다. 하지만 과분할 영역의 결과는 차후 깊이 값을 부여하는 과정에서 오차를 발생하게 하므로 최소 화소 이하의 영역 처리를 수행한다. 이웃한 영역들과의 유사도를 비교하기 위해 식 (2) 와 같이 비용합수를 계산한다.

$$C(R_i, R_j) = C_r(R_i, R_j) + C_u(R_i, R_j) + C_v(R_i, R_j) \quad (2)$$

각각의 영역은 주위 영역들과의 색상의 유사도를 비교하여, 가장 작은 비용값을 가진 영역과 융합하는 과정을 정해진 최소의 화소값 이상이 될 때까지 반복한다.

나. 프레임 그룹 설정

영상의 움직임 정보를 추정하기 위하여 일반적으로 전,후 프레임을 이용한 움직임 추정 연산을 수행하지만, 전후 프레임의 움직임 정보가 거의 존재하지 않는 영상에 의한 잘못된 결과값으로 추후 3차원 변환 과정에서의 오차가 발생하게 된다. 이를 방지하기 위해 연산할 프레임을 그룹화 하여 양방향 탐색을 수행토록 하였다.

프레임을 그룹화 하기 위해 장면전환점 검출연산을 수행하게 되는데, 이는 다음 식 (3)과 같다.

$$\sum_{j=1}^N |H_i(j) - H_{i+1}(j)| > T \quad (3)$$

$H_i(j)$ 는 i 번째 영상에서의 히스토그램 값 j (0 ~ 255)의 누적 히스토그램 분포를 나타내고 T는 허용 오차값으로 전체 화소의 개수의 30%로 설정하였다. 히스토그램 분포의 차이가 허용오차 이상이라면 장면 전환점이라고 판단한다. 장면 전환점이라고 판단된 i 번째 영상은 그룹의 마지막 영상으로 설정되고, 설정된 시작 영상부터 마지막 영상까지의 그룹화된 입력 영상들로부터 움직임 추정 연산을 수행하게 된다.

다. 깊이 정보 생성



그림 1. 움직임 정보의 유무에 따른 탐색 영상 변경

그림 1과 같이 i번째 프레임과 i+1번째 프레임간의 움직임 추정 연산 수행 중 구해진 움직임 값 (intensity)이 영상의 전체 화소 개수의 10% 이하로 존재한다면 움직임 추정을 수행 하지 못했다고 판단하고, i번째 영상은 i+2 번째 영상과 움직임 추정 연산을 수행하게 한다. 마찬가지로 i+2 영상과의 움직임 추정을 수행하지 못했다고 판단한다면 다음 프레임으로 연산을 넘겨준다. 설정된 마지막 영상, 혹은 연산할 영상 그룹이 마지막 영상과 인접하여 수행 가능한 영상그룹이 존재하지 않는다면, 영상 그룹의 탐색을 역방향으로 수행한다. i번째 영상과 i+1 번째 영상 혹은 그 이후의 영상에서의 움직임 정보가 존재하지 않는다고 판단되면, i-1번째 영상방향으로의 탐색 및 연산을 수행한다.

영상의 움직임 정보를 얻기 위해 다음 식을 통하여 움직임 추정 연산을 수행한다. 움직임 추정은 SAD (sum of absolute difference) 값을 비용값으로 사용하고 spiral search 후 결과값이 허용오차 이하일 때 중지한다.

구해진 움직임 값을 깊이 정보로 변환하기 위하여 식 (4)의 과정을 거친다.

$$D(x, y) = Motion_x + Motion_y \quad (4)$$

여기에서 $D(x, y)$ 는 (x, y)좌표 화소에서의 깊이값을 의미한다. $Motion_x$ 는 x 방향으로의 움직임 크기(intensity)이고, $Motion_y$ 는 y 방향으로의 움직임 크기이다.

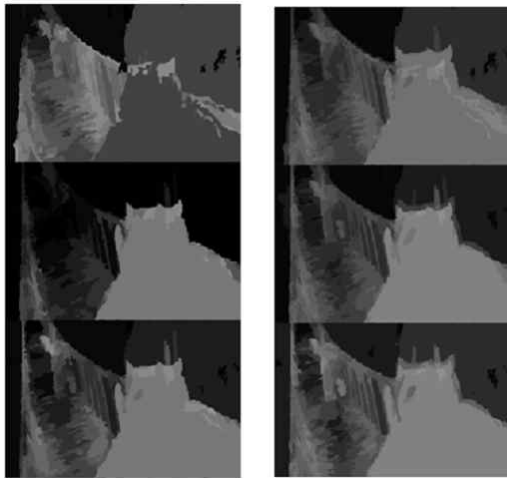
$D(x, y)$ 는 많은 오차값을 가지고 있고, 한 영역 안에서 화소들 간의 편차가 크게 존재하므로 다음과 같은 처리가 필요하다. 구한 $D(x, y)$ 에 대하여 영역 내의 화소들의 깊이 값 평균을 구해서 한 영역내에서 같은 값을 가지도록 한다. 이 과정을 식 (5)에 나타내었다.

$$D(R_i) = \frac{1}{m \times n} \sum_{k=0}^{m \times n - 1} D(k) \quad (5)$$

위의 식에서 $D(R_i)$ 는 $m \times n$ 화소를 갖는 i 번째 영역에서의 깊이 값을 의미하며 $D(k)$ 는 영역안에서의 k 번째 화소의 깊이 값을 의미한다.

라. 후처리

생성된 깊이맵은 후처리 기법을 사용함으로써 보다 신뢰도 높은 결과를 얻을 수 있다. 영상 전체에서 각각의 프레임은 전후프레임간의 차이가 발생하여 오차를 발생시킨다. 앞뒤 프레임 혹은 사용자가 정한 프레임 개수만큼의 평균값 처리를 하여 보다 안정적인 깊이맵을 구할 수 있다. 다음 그림 2에 그 결과를 나타내었다.



(a) (b)

그림 2. 평균값 처리 결과

(a) 평균값 처리 전 영상 (b) 평균값 처리 후 영상

3. 결론

본 논문에서는 영역분할과 움직임 정보를 이용한 깊이맵 생성 기법을 제안하였다. 영역을 분할하고, 움직임 정보를 이용하여 분할된 영역에 깊이를 부여하는 과정에서 여러 변수들에 의해 많은 오차가 발생하게 된다. 그 중 부정확한 움직임 정보로 인한 오차는 깊이맵을 제대로 생성할 수 없게 하고, 생성된다 하더라도 신뢰도가 매우 낮은 깊이맵이므로 추후의 3D 변환 과정에서 사용할 수 없다. 이러한 문제점을 해결 하기 위해 본 논문에서는 영상을 그룹화 하고 움직임정보의 존재 유무를 판단하여 판단기준에 어긋나면 새로운 연산을 수행토록 하였다. 움직임 정보가 존재하지 않았던 영상에서 잘못된 깊이맵이 부여되던 문제점을 해결하였고, 움직임정보의 신뢰도를 향상시켰다. 초기 깊이맵이 생성되면, 전후 프레임 혹은 사용자가 정한 더 많은 수의 프레임과의 평균값을 부여하는 과정을 수행하므로써 영상간의 깊이맵 차이를 줄일 수 있었고, 최종적으로 신뢰도 높은 깊이맵 영상이 생성되는 것을 확인할 수 있었다..

<감사의 글>

본 연구는 서울시 산학연 협력사업(NT080528)과 지식경제부 및 한국산업기술평가관리원의 IT산업원천기술개발사업[KI002058], 대화형 디지털 홀로그램 통합서비스 시스템의 구현을 위한 신호처리 요소 기술 및 SoC 개발사업의 일환으로 수행되었음.

참 고 문 헌

- [1] 백윤기, 최미남, 박세환, 유지상, "2차원 동영상으로부터 다시점 동영상 생성 기법", 한국통신학회논문지, Jan. 2008, Vol.33 No. 1 pp.53-61.
- [2] 홍호기, 백윤기, 이승현, 김동욱, 유지상, "2D H.264 동영상의 3D 입체 변환", 한국통신학회논문지, 제31권 제12C호, pp.1208-1215,