영역분할과 컬러 특징을 이용한 건물 인식기법

Building Recognition using Image Segmentation and Color Features

허 정 훈¹, 이 민 철[†]

Heo Jung-Hun¹, Lee Min-Cheol[†]

Abstract This paper proposes a building recognition algorithm using watershed image segmentation algorithm and integrated region matching (IRM). To recognize a building, a preprocessing algorithm which is using Gaussian filter to remove noise and using canny edge extraction algorithm to extract edges is applied to input building image. First, images are segmented by watershed algorithm. Next, a region adjacency graph (RAG) based on the information of segmented regions is created. And then similar and small regions are merged. Second, a color distribution feature of each region is extracted. Finally, similar building images are obtained and ranked. The building recognition algorithm was evaluated by experiment. It is verified that the result from the proposed method is superior to color histogram matching based results.

Keywords: building recognition, watershed algorithm, region based image retrieval, RAG(Region Adjacency graph), IRM(Integrated Region Matching)

1. 서 론

도시 환경에서 주변의 사물을 인식하는 것은 다양한 비전 응용 분야에 적용될 수 있는 중요한 기술이다. 이중 건물과 같은 인공구조물은 주변환경에 대한 정보 및 특정 장소를 대표하는 랜드마크(landmark)로서 활용 될 수 있으므로 많은 분야에서 이를 활용하기 위한 연구가 증가하고 있다. 특히 주변환경을 인식하여 그 환경에 대한 부가정보를 보여주는 증강현실(augmented reality) 분야와 위치인식 및 주행에 관련하여 많은 연구가 진행되고 있다. 증강현실과 관련한 연구로 Fritz는 모바일 폰의 카메라와 GPS 정보를 이용한 건물인식을 통해 관련된 정보를 표시해주

기 위한 연구를 수행하였다^[1]. 한편 위치인식과 관련한 연구로 Lee는 주변에 많은 고층건물로 인해 GPS 신호가 제대로 수신되지 않아 정확한 위치정보를 얻기 어려운 도심 영역, 소위 “urban canyons” 라 부르는 영역에서 건물인식을 통하여 GPS 정보를 보정하고자 하였다^[2]. 또한 Schindler와 Szeliski는 구글(Google map)의 스트리트뷰(street view)와 같은 GPS정보가 포함된 영상 데이터베이스들과 현재 위치에서 획득한 건물 영상의 정합을 통해 현재 위치를 추정하는(vision-based localization) 방법에 대해 연구하였다^[3].

이런 건물 인식을 위한 방법으로서 SIFT(Scale Invariant Feature Transform)^[4]와 같은 영상특징을 이용한 방법과 건물의 기하학적 특징 또는 건물의 색상, 질감, 형상을 이용하는 내용기반영상검색(Content-Based Image Retrieval)을 들 수 있다. SIFT 기반 알고리즘은 객체의 모서리나 꼭지점 등에서 생성되는 특징들을 이용하여 영상간 정합을 하는데, 영상의 크기 및 회전, 조명등의 변화에도 뛰어난 정합 성능을 보여 건물인식뿐만 아니라 다양한 객체인식 문제에 많이 사용되고 있다^[5,6]. 하지만 많은 계산을 필요로 하며

Received : Nov. 5. 2012; Reviewed : Dec. 6. 2012; Accepted : Apr. 30. 2013

※ This research was supported by MKE(The Ministry of knowledge Economy), Korea, under the Human Resource Development Program for Specialized Navigation/Localization Technology Research Center support program supervised by the NIPA(National IT Industry Promotion Agency (NIPA-2012-H1502-12-1002))

¹ Interdisciplinary program in robotics, Pusan National University, Jangjeon 2-Dong, Geumjeong-Gu, Busan, Korea (gasmaster2@naver.com)

[†] Corresponding author: Mechanical Engineering, Pusan National University, Jangjeon 2-Dong, Geumjeong-Gu, Busan, Korea (mclee@pusan.ac.kr)

특징점이 많이 추출되지 않는 영상은 인식이 어렵다는 단점이 있다. 건물의 색상, 질감, 형상을 이용하는 방법은 내용기반 검색 (CBIR: Content-Based Image Retrieval) 을 기반으로 한다. CBIR은 문서위주의 자료에서 최근 급속히 이미지 등의 자료가 늘어남에 따라 이를 텍스트 기반이 아닌 사용자가 제시하는 영상으로 검색하기 위하여 영상에서 색상, 질감, 형상과 같은 값들을 이용하여 유사한 영상을 찾고자 하는 연구분야이다^[7-10]. 이중 컬러 히스토그램을 이용한 방법은 건물 인식분야에서 상대적으로 좋은 성능을 보여준다^[11]. 이는 건물의 색상 구성이 그 건물을 표현하는 좋은 특징이 될 수 있음을 보여준다. 하지만 실제 실외환경에서 건물 촬영 시 주변에 많은 객체들이 존재하며, 상대적으로 건물이 차지하는 영역이 적을 수록 건물 영상의 컬러 히스토그램 변화가 크기 때문에 그 인식률이 떨어지게 된다. Fig. 1(a)는 이러한 상황을 보여준다.

본 논문에서는 이런 단점을 보완하기 위한 방법으로서 Fig. 1 (b)에서 볼 수 있듯이 영역분할 후 각 영역의 컬러 특징을 이용한 건물 인식 기법을 제안한다.

제시하는 건물 인식 기법은 다음과 같다. 입력영상 watershed 알고리즘을 이용하여 영역을 분할하고, 그 연결성과 컬러 특성, 그리고 영역간의 크기 및 관계를 고려하여 영역을 병합한다. 최종 분할 영역 중 하늘, 나무와 같은 영역은 컬러모델을 이용해서 제거하며, 다음 각 영역

(b) comparison of color histogram through image segmentation
Fig. 1. Comparison of between conventional color histogram and proposed method

Fig. 2. Flow chart of proposed method

에서 색상과 채도 값에 관한 히스토그램을 추출한다. 이를 바탕으로 질의 영상과 DB에 저장된 영상들의 각 영역간에 대해서 거리행렬을 구하고 각 영역크기에 따른 가중치를 적용하여 비교 영상들의 유사도 값을 구한다. 이 값을 이용하여 주어진 건물 영상과 가장 유사한 영상을 구하게 된다. Fig. 2는 전술한 과정에 대한 순서도를 보여준다.

2. 영역 분할 및 병합

2.1 Watershed 알고리즘을 이용한 영역 분할

영상분할은 크게 경계선(edge)을 기반으로 한 방법과 영역(region)을 기반으로 한 방법으로 나뉘어 지는데 경계선에 기반한 방법의 경우 경계선을 바탕으로 분할하므로 잡음에 민감하고 폐곡선 형태의 경계선을 얻기 어렵다. 반면 watershed의 경우 영역간 명암도의 유사성에 근거하는 방법으로 잡음이 많은 환경에서 물체를 잘 분할하는 편이지만 물체간의 세부적인 경계선을 명확하게 추출하기 어려운 단점이 있다. 하지만 다른 영역분할 알고리즘에 비해 특정 파라미터를 필요로 하지 않으며 전처리 및 후처리 과정을 적용하기 쉽다. 또한 최근에는 영상의 경계선 정보를 이용하여 외곽선 정보를 추출하고 이를 이용하여 watershed 분할을 수행 함으로서 영역기반으로 영역분할시의 단점을 보완하는 방향으로 연구가 이루어지고 있다^[12]. 본 논문에서는 건물의 경계선을 유지하면서 영역분할을 하기 위해서 앞서 설명했듯이 경계선을 추출한 후 거리

Fig. 3. Image segmentation process and its results

변환을 적용하여 watershed 알고리즘을 적용하였다. 이를 이용한 영역분할 과정은 다음과 같다.

우선 가우시안 필터(Gaussian filter)를 사용하여 잡음을 제거하고, Canny 경계검출자에 의해서 경계지도(edge map)을 얻은 후 이를 거리변환 하여 watershed 변환을 수행한다. Fig. 3은 전술한 영역분할 과정 및 그 변환 결과를 보여준다.

2.2 영역 병합 과정

2.2.1 영역 인접 그래프를 이용한 병합

Fig. 3(d)에서 보듯이 watershed 알고리즘의 대표적 단점인 과분할(over-segmentation)현상을 해결하기 위해 세분화된 영역을 의미 있는 영역으로 병합하는 과정이 필요하다. 이를 위해 영역 인접 그래프(RAG: Region Adjacency Graph)를 사용한다^[13]. 영역 인접 그래프는 서로 인접한 영역간의 관계를 그래프로 표현하는 방법이다. 영역들의 정보를 정점으로 하며 두 영역 사이의 관계를 간선으로 표현하는데 여기서 두 영역 사이의 관계는 8 이웃관계로 인접하는 픽셀의 개수로 정의하였다. 아래 Fig. 4(a)는 4×4 영상이며 이를 앞서 정의한 영역 인접 그래프로 표현하면 Fig.

Fig. 4. 4×4 label image and its region adjacency graph

4(b)와 같이 표현될 수 있다.

영역 인접 그래프를 이용하여 과분할된 영역을 제거하기 위해서 영상에서 두 영역간의 경계가 실제 각 영역의 경계인지 과분할 된 영역인지 판단하는 식 R_c 를 정의한다. 두 영역의 인접픽셀 개수를 N_p 라 하고 이 중 edge 성분에 해당하는 픽셀의 개수를 N_e 라 정의한다.

$$R_c = \frac{N_e}{N_p} \quad (1)$$

식 (1)을 이용하여 값을 계산하면 R_c 값이 0에 가까울수록 두 영역의 경계가 과분할 된 영역의 경계임을 나타내며 1에 가까울수록 실제 영역경계에 해당함을 나타낸다. 영상내의 모든 경계에 대하여 R_c 가 0.2 이하는 서로 병합하도록 한다. R_c 의 값은 실험적으로 선정하였으며 한 영역이 인접한 영역으로 병합될 시 아래 Fig. 5와 같은 순서에 의하여 병합된다. 여기서 S영역은 병합대상이 되는 영역이며 이 S영역이 접하고 있는 두 영역 L과 K중 L에 병

Fig. 5. Region merging process using region adjacency graph

Fig. 6. Merging results using region adjacency graph

Fig. 7. Merging results using color histogram

합된다고 가정하면 Fig. 5(a)는 L과 K 두 영역이 서로 인접해 있는 경우의 병합 과정을 나타내며 Fig. 5(b)는 L과 K 두 영역이 서로 인접하지 않는 경우의 과정을 보여준다. Fig. 6은 임의의 한 영역이 인접한 영역으로 병합하기 전과 병합후의 결과를 나타낸다.

2.2.2 컬러 히스토그램 유사도를 이용한 병합

건물 영역 내부와 같은 영역은 비교적 밝기변화가 적으며 따라서 영역인접 그래프와 경계선 정보를 이용한 병합 과정에서 많은 영역들이 병합된다. 하지만 지역적으로 생기는 밝기차이(그림자 및 광량의 차이)로 인해 병합되지 않는 영역이 발생한다. 이런 경우 영역 경계에서의 변화가 아닌 두 영역내부의 컬러 분포 유사도를 이용하여 병합을 수행한다.

이를 위해서 컬러 히스토그램을 사용하였다. 비교하고자 하는 각 영역들의 컬러 히스토그램을 만든 후 영역인접 그래프를 이용하여 서로 인접한 영역간의 컬러 히스토그램 유사도를 비교하여 영역을 병합하였다. 컬러 공간은 HSV 색 공간을 이용하였으며 색상과 채도를 각각 36 bin과 20 bin으로 분할하여 각 영역별로 2 차원 히스토그램을 생성하였다. 히스토그램의 유사도를 계산하기 위한 거리 척도로 식(2)의 히스토그램 교차 값을 이용하였다.

$$d_{intersection}(H_1, H_2) = \sum_i \min(H_1(i), H_2(i)) \quad (2)$$

Fig. 7은 영역들의 컬러 히스토그램과 그 인접성을 이용하여 병합한 결과이다.

2.2.3 크기 및 영역 관계를 이용한 병합

Fig. 7(b)의 병합결과에서 보여지듯이 하늘, 건물, 그리고 도로와 같이 유사한 색상을 가지며 그 경계가 뚜렷한 영역들은 비교적 큰 영역들로 병합되었음을 알 수 있다. 하

지만 차량, 나무 등의 작은 객체들은 병합되지 않고 남아있음을 알 수 있다. 또한 건물 영역내부에서는 창문 및 건물내의 색상 차이로 인해 내부에 포함된 작은 영역들이 병합되지 않고 남아있게 된다. 이를 병합하기 다음의 세가지 조건을 적용하여 작은 영역들을 병합한다.

case 1: 일정 크기 이하의 영역들 중 서로 인접하여 뭉쳐 있는 경우 병합

case 2: 큰 영역 내부에 포함되거나 둘러 싸여 있는 영역들은 병합

case 3: 상대적으로 큰 영역들의 경계에 존재하는 작은 영역들을 더 많은 픽셀을 이웃 하고 있는 영역으로 병합

Fig. 8은 위 3가지 조건에 해당하는 영역들을 보여준다. 대부분의 영상에서 컬러 유사도를 이용한 병합 후에는 상대적으로 큰 영역들은 영상의 대부분 영역을 차지하며 나머지 영역들은 0 ~ 1% 사이의 작은 영역들만 남게 되는데 이를 이용하여 일정 크기 이하의 영역들의 조건을 영상 전체크기에서 0.05% 이하가 되는 영역으로 선정하였다. Fig. 9는 크기 및 영역간 관계를 이용한 병합 결과를 보여준다.

Fig. 8. Classification of small regions

Fig. 9. Merging results using region size and relationship

3. 컬러 특징 추출 및 영상 매칭 과정

3.1 하늘 및 나무 영역 제거

영역 분할 후 하늘, 나무, 풀 등과 같이 건물과 관련이 없는 영역은 제거하도록 한다. 이를 위해서 하늘과 나무의 컬러 모델을 선정하고 컬러 모델과 유사한 영역을 제거하였다. 하늘 영역의 경우 흐린 날, 맑은 날로 크게 분류할 수 있다. 건물 영상 중 흐린날과 맑은 날의 하늘이 포함된 영상들 100개씩을 선정하여 HSV색공간에서의 평균값과 분포를 이용하여 컬러 모델을 선정하였다. 나무의 경우 짙은 녹색계열의 색상을 가지며 채도 값이 낮게 나타나는 침엽수 계열의 나무 영상들에 한정하여 컬러 모델을 선정하였다. Table 1은 그 결과를 보여준다. Fig. 10은 분할된 영역 중 제시한 컬러 모델을 통하여 하늘 및 나무 영역을 제거한 결과이다.

Table 1. Values of color models selected

regions	hue	saturation	value
clear sky regions	180 - 210	0 - 255	180 - 255
cloudy sky regions	0 - 360	0 - 50	220 - 255
tree regions (coniferous)	70 - 120	50 - 200	0 - 52

(a) cloudy sky regions (b) clean sky regions (c) tree regions
Fig. 10. removed regions (black regions) using selected color models

3.2 IRM(Integrated Region Matching) 개요

분할 된 영상의 유사도를 측정하기 위한 방법으로 영역 기반 검색 알고리즘 IRM(Integrated Region Matching)을 사용하였다^[14]. 이는 기존의 분할된 영역단위의 검색 방식에 비해 Fig. 11과 같이 한 영상 내에서 분할된 여러 영역을

Fig. 11. Comparison of between conventional region-based matching and IRM matching

영역집합으로 묶고 이들간의 비교를 통해 영상간의 유사도를 판단하므로 영역분할이 부정확한 경우에도 비교적 정확한 결과를 얻을 수 있다.

IRM의 과정은 다음과 같다. 두 영상에서 분할된 영역들의 집합을 각각 $R_1 = \{r_1, r_2, \dots, r_m\}$ 과 $R_2 = \{r'_1, r'_2, \dots, r'_n\}$ 로 가정한다. 여기서 r_i 와 r'_j 는 각 영상에서 분할된 영역을 나타낸다. 두 영역 r_i 와 r'_j 간의 거리(inter-region distance)는 $d(r_i, r'_j)$ 로 정의되며 줄여서 $d_{i,j}$ 로 표현한다. 거리 평가함수는 다음 3.3 절에서 설명하도록 한다. 두 영상에 대한 거리(image-to-image similarity) $d(R_1, R_2)$ 는

$$d(R_1, R_2) = \sum s_{i,j} d_{i,j} \quad (3)$$

식 (3)과 같이 정의되며 여기서 $s_{i,j}$ 는 중요도(significance credit)행렬로서 두 영역 r_i 와 r'_j 의 매칭 시 중요도에 따라서 가중치가 할당된다. 여기서 $s_{i,j} \geq 0$ 이며

$$S = \begin{bmatrix} s_{1,1} & s_{1,2} & \dots & s_{1,n} \\ s_{2,1} & s_{2,2} & \dots & s_{2,n} \\ \dots & \dots & \dots & \dots \\ s_{m,1} & s_{m,2} & \dots & s_{m,n} \end{bmatrix} \quad (4)$$

위 행렬 (4)와 같이 구성된다. 중요도 행렬의 값을 할당하기 위해서 식 (5),(6)의 조건을 만족하도록 값을 할당한다. 여기서 p_i 는 영상1의 영역 r_i 에 대한 중요도 행렬의 합을 의미하며 p'_j 는 영상2의 영역 r'_j 에 대한 중요도 행렬의 합을 의미한다.

$$\sum_{j=1}^n s_{i,j} = p_i, i = 1, \dots, m \quad (5)$$

$$\sum_{i=1}^m s_{i,j} = p'_j, j = 1, \dots, n \quad (6)$$

정규화를 위해서 $\sum_{i=1}^m p_i = \sum_{j=1}^n p'_j = 1$ 이 되도록 하여야 하며, 이를 위해 다음과 같은 과정1~8 에 의해서 중요도 행렬의 값 $s_{i,j}$ 가 할당된다.

- 1: $L = \{ \}$, 과 $M = \{(i,j) \mid i = 1, \dots, m; j = 1, \dots, n\}$ 을 설정한다.
- 2: $d_{i,j}$ 가 최소가 되는 $(i,j) \in M - L$ 에 대해 (i,j) 를 (i',j') 로 둔다.
- 3: $\min(p_{i'}, p'_{j'}) \rightarrow s_{i',j'}$.
- 4: 만약, $p_{i'} < p'_{j'}$ 이면 $s_{i',j} = 0$ 으로 두고 그렇지 않다면 $s_{i,j'} = 0, i \neq i'$
- 5: $p_{i'} - \min(p_{i'}, p'_{j'}) \rightarrow p_{i'}$
- 6: $p'_{j'} - \min(p_{i'}, p'_{j'}) \rightarrow p'_{j'}$
- 7: $L + \{(i',j')\} \rightarrow L$
- 8: 만약, $\sum_{i=1}^m p_i > 0$ and $\sum_{j=1}^n p'_j > 0$ 이면 2번째 단계로 가서 반복, 그렇지 않다면 종료한다.

여기서 p_i 와 p'_j 의 값은 각 영역의 중요도에 따라서 할당이 될 수 있는데 본 논문에서는 영역이 전체 영상에서 차지하는 비율로 선정하였다. 과정 1~ 8을 설명하기 위하여 Fig. 12와 같이 두 영역으로 분할된 두 영상 R_1 과 R_2 를 예로 들어 설명하고자 한다.

주어진 두 영상의 각 영역간의 거리 $d(r_i, r'_j)$ 를 구하여 식(7)과 같은 거리행렬 $\{d_{i,j}\}$ 를 얻는다고 가정한다.

$$\{d_{i,j}\} = \begin{pmatrix} 0.2 & 0.7 \\ 0.4 & 0.9 \end{pmatrix} \quad i = 1, 2, j = 1, 2 \quad (7)$$

영상 R_1 의 두 영역 r_1, r_2 가 전체 영상영역에서 차지하는 비중이 각각 40% 와 60%라고 가정하면 $p_1 = 0.4, p_2 = 0.6$ 로 할당된다. 영상 R_2 에 대해서도 각 영역의 크기에 따라 $p'_1 = 0.2, p'_2 = 0.8$ 이 된다고 가정한다. 앞서 설명한 과정 1에서와 같이 $L = \{ \}$ $M = \{(1,1), (1,2), (2,1), (2,2)\}$ 로 할당되며 여기서 $d_{i,j}$ 중

Fig. 12. Example of two segmented images

최소값은 0.2 이므로 $(i',j') = (1,1)$ 이 되며 과정3에 의해 $\min(p_1, p'_1) \rightarrow s_{1,1}$ 이고 $p_1 = 0.4, p'_1 = 0.2$ 이므로 $s_{1,1} = 0.2$ 가 되며 $p_1 > p'_1$ 이므로 $s_{2,1} = 0$ 이 할당된다. 과정 5~7에 따라 $p_1 = 0.2, p'_1 = 0$ 이 되며 $L = \{(1,1)\}, M = \{(1,2), (2,1), (2,2)\}$ 이 된다.

과정 8의 조건을 만족하므로 M 의 값중 $d_{i,j}$ 가 최소가 되는 값 $(i',j') = (2,1)$ 에 대해 과정2부터 반복 수행한다. 모든 반복수행이 끝나게 되면 식(8)과 같은 중요도행렬을 얻을 수 있다.

$$S = \begin{pmatrix} 0.2 & 0.2 \\ 0 & 0.6 \end{pmatrix} \quad (8)$$

따라서 식(3)에 의하여 두 영상간의 유사도 값은 0.72임을 구할 수 있다.

3.3 거리 평가함수

각 영역의 컬러분포 유사도를 평가하기 위하여 영상을 빛의 변화에 강인한 HSV 색공간으로 변환하고 각 영역별로 색상(hue)을 16등분, 채도(saturation)을 8등분, 명도(value)값을 4등분한 3차원 컬러 히스토그램을 추출하였다. 이 히스토그램들간의 유사도를 평가하기 위한 함수 $d(r_i, r'_j)$ 로서 카이-제곱(chi-square) 거리를 사용하였다. 카이-제곱 거리는 두 히스토그램의 유사도를 측정하기 위한 대표적인 방법 중 하나로서 식(9)와 같이 정의된다^[5]. 여기서 H_1, H_2 는 비교대상이 되는 두 히스토그램을 의미하며 i 는 히스토그램에서 bin의 개수를 의미한다.

$$d_{chi-square}(H_1, H_2) = \sum_i \frac{(H_1(i) - H_2(i))^2}{H_1(i)} \quad (9)$$

4. 실험 결과

4.1 ZubuD 건물 database

제안한 방법의 성능 평가를 위해서 스위스 연방 공과 대학에서 만든 “Zurich Buildings Database for Image Based Recognition”(ZubuD)를 사용하였다^[16].

취리히 시내의 201 개의 건물에 대하여 각각 다른 시점에서 촬영된 영상 5개씩 총 1005개로 구성된 트레이닝 영상과 115개의 질의 영상으로서 구성되어 있다.

4.2 실험 방법 및 결과

실험을 위하여 건물 영상들을 크게 3가지로 분류하였다. ZubuD 영상들은 Fig. 13(a) 와 같이 영상에서 건물이 차지하는 영역이 크며 그 시점변화도 적은 건물 영상들과 Fig. 13(b)와 같이 건물 영역 및 그 시점변화가 상대적으로 큰 영상들과 같이 2가지로 분류하였다. Fig. 13(c)는 실제 환경에서 촬영한 영상들을 보여준다. Fig. 13(c)를 보면 알 수 있듯이 건물의 특징이 잘 나타나도록 촬영된 ZubuD 영상들에 비해 실제 환경에서 얻은 영상은 건물 영역이 차지하는 크기가 크지 않으며 촬영 시점변화도 크며 또한 건물 주변의 자동차, 나무, 도로, 사람 등 많은 객체들이 같이 포함되는 복잡한 영상인 경우가 많다. 따라서 이런 영상에 대하여 본 논문에서 제안한 방법이 더 좋은 성능을 보임을 실험을 통해 확인하고자 한다. Fig. 13. (c)의 영상은 부산대 건물들을 대상으로 촬영하였으며, 30 개의 건물에 대하여 각 5개씩 다른 시점에서 촬영된 영상 150개로 구성되어 있다. 같은 방식으로 Fig. 13. (a)-(b) 또한 ZubuD 영상 중 건물 30개에 대하여 다른 시점으로 촬영된 영상 150개씩을 선정하였다.

이렇게 선정된 (a)-(c)의 건물 영상들에 대하여 제안한 방법과 기존의 컬러 히스토그램방법을 사용하여 건물인식 성능을 비교 평가하고자 한다. Fig. 14는 주어진 건물 영상

Fig. 13. Classification of building images

Fig. 14. Query image and its corresponding building images

(a)에 대하여 제안한 알고리즘을 이용하여 전체 영상 데이터베이스에 대하여 유사도를 계산하고 그 유사도가 높은 순으로 좌측부터 나열한 영상이다. (b)의 각 그림 상단의 소수값은 계산된 유사도 값을 의미한다. IRM을 통하여 계산된 유사도 값은 0~1의 값을 가지며 유사한 건물 영상일수록 0에 가까우며 반대일수록 1에 가까운 값을 가진다. 아래의 zubuD-114-3 은 ZubuD 영상중 114번 건물의 3번째 영상을 의미한다. 결과에서 2번째영상은 잘못된 결과임을 알 수 있는데 이를 정량적으로 평가하기 위하여 정확도(precision)와 재현율(recall)을 이용한다. 두 식의 정의는 식 (10), (11)과 같다^[7].

$$P = \frac{\text{검색된 결과 영상중 관련 영상의 개수}}{\text{검색결과 영상의 개수}} \quad (10)$$

$$R = \frac{\text{검색된 결과 영상중 관련 영상의 개수}}{\text{관련 영상의 개수}} \quad (11)$$

정확도와 재현율은 대부분의 경우 서로 반비례하는 관계를 가진다. Fig. 14와 같은 경우 검색 결과 영상으로서 유사도가 가장 높은 첫번째 영상만 이용한다면 정확도는 1/1(100%)가 되며 재현율은 1/5(20%)가 된다. 반면 가장 유사도가 떨어지는 8번째 영상까지 이용한다면 정확도는 5/8(63%)가 되며 재현율은 5/5(100%)가 된다. 평균정확도 (AP: average precision)는 식(12)와 같이 정의되며 N_r 은 전체 영상 중 주어진 영상과 관련된 영상의 수를 의미하며 R_n 은 관련 영상의 개수가 n 개의 경우 재현율을 의미한다.

$$AP = \frac{1}{N_R} \sum_{n=1}^{N_R} P_q(R_n) \quad (12)$$

Fig. 14의 경우 AP는 $(1/1 + 2/3 + 3/4 + 4/5 + 5/8)/5 = 0.77$ 임을 알 수 있다. 전체 영상에 대한 성능을 평가하기 위하여 식 (13)과 같은 평균정확도의 평균값(MAP: Mean Average Precision)을 사용하였다. 여기서 Q 는 전체 질의영상(제시 영상)의 집합을 의미하며 $|Q|$ 는 질의영상의 개수를 의미한다.

$$MAP = \frac{1}{|Q|} \sum_{q \in Q} AP(q) \quad (13)$$

앞서 제시한 Fig. 13(a)~(b)에 대하여 MAP를 계산할 결과는 Table 2와 같다. 결과에서 보듯이 Fig. 13(c)의 경우와 같이 주변에 많은 객체가 존재하며 큰 시점변화를 가지는 경우에는 제안한 방법이 확실히 더 좋은 결과를 보임을 확인할 수 있다.

Fig. 13(b)의 경우에도 영역분할을 통한 방법이 더 좋은 결과를 보임을 알 수 있다. 이는 Fig. 15에서도 보듯이 히스토그램의 변화가 큰 경우 제안한 방법이 더 효율적임을 알 수 있다. 반면 Fig. 13(a)와 같이 히스토그램의 변화가

Table 2. MAP of each building DB images

	[Fig 13](a)	[Fig 13](b)	[Fig 13](c)
Proposed method	72.1[%]	74.5[%]	48.3[%]
Color histogram	78.5[%]	72.5[%]	36.2[%]

Fig. 15. Change of histogram shape according to building images

(a) query image and its segmented image

(b) result of color histogram method

(c) result of proposed method

Fig. 16. Result caused by poor image segmentation

없는 경우는 오히려 떨어지는 성능을 보이는데 이는 Fig. 16(a)에서와 같이 분할이 잘못된 경우를 주된 이유로 들 수 있다.

그 결과로서 Fig. 16(b)~(c)를 보면 알 수 있듯이 좋지 못한 영역분할 결과로 인해 기존 방법에 비해 더 좋지 못한 결과를 얻음을 볼 수 있다. 그 이유는 Fig. 17의 경우를 예로 들 수 있다. 상단의 분할 영상은 영상의 대부분을 차지하는 1번 영역과 2번 영역이 잘 분할되었음을 알 수 있다. 또한 하늘에 해당하는 영역도 정확하게 제거 되었음을 확

Fig. 17. Example of bad results in building recognition

인할 수 있다. 반면 아래 건물 영상은 하늘 영역이 전혀 제거 되지 못했으며 오히려 건물 영역을 잘 못 제거 하였음을 알 수 있다. 또한 두 건물 영상은 영상 전체로는 유사하지만 분할된 영상영역 단위로는 그 유사성이 떨어짐을 알 수 있다.

Fig. 16도 이런 경우로 볼 수 있다. 이런 이유로 Table 2 의 Fig. 13(a)가 제안한 방법에 비해 더 좋은 결과를 보임을 알 수 있다.

5. 결 론

본 논문에서는 건물의 컬러 특징을 이용하여 건물을 인식하기 위해서 watershed 영역 분할 알고리즘을 이용하여 영상을 분할하고, 분할된 영역들간의 연결성, 색상유사도 그리고 영역의 크기를 이용하여 영역 병합을 함으로써 영상을 건물, 하늘, 도로 등과 같이 의미 있는 영역으로 분할하였다. 이중 하늘, 나무와 같이 여러 영상에서 나타나는 공통적인 영역을 제거하기 위하여 이에 해당하는 HSV 컬러 모델을 선정하였다. 이를 이용하여 영상에서 이에 해당하는 영역을 제거하여 이로 인한 영향을 최소화 하도록 하였다. 또한 분할된 영역들이 어느 정도 부정확하더라도 IRM을 이용하여 그 유사도를 비교하게 함으로써 좋은 인식율을 보일 수 있도록 하였다. 하지만 영역분할의 결과가 건물, 도로, 하늘 등과 같이 명확하게 구분이 되지 않는 경우 그 인식율이 떨어짐을 확인 할 수 있었다. 또한 실제 영상에서는 그 인식율이 크게 떨어짐을 확인할 수 있었다. 이를 보완하기 위해서 향후 연구 과제로서 영역 분할의 정확성을 더욱 향상 시키기 위하여 분할 및 병합 과정에서 건물의 질감(texture), 기하학적 특징(소실점, 반복구조)등을 활용한 연구가 필요하다. 또한 전체 영상의 컬러 분포(히스토그램)를 분석하여 주어진 건물영상 특성에 따라 적절한 분할 영역 개수를 선정하는 부분과 인식율을 향상 시키기 위하여 컬러 특징 뿐만 아니라 질감, 형상 정보도 같이 이용하여 그 유사도를 계산하는 연구가 수행 되어야 할 것이다.

References

- [1] G. Fritz, C. Seifert, and L. Paletta, "A Mobile Vision System for Urban Detection with Informative Local Descriptors," *Proc. of the 4th Conf. Computer Vision Systems*, pp. 30, Jan 2006.
- [2] Lee Chia-Hsiang, Su, Yu-Chi, Chen, Liang-Gee "Accurate Positioning System based on Street View Recognition," *IEEE Int. Conf. Acoustics, Speech and Signal Processing(ICASSP)*, pp.2305-2308, 2012.
- [3] Brown M. Schindler G, Szeliski R. "City-Scale Location Recognition," *IEEE Int. Conf. Computer Vision and Pattern Recognition(CVPR)*, pp.1-7, 2007.
- [4] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110, Jan 2004.
- [5] Zach C, Frahm J.-M., Bischof H. "From Structure-from-motion point clouds to fast location recognition," *IEEE Int. Conf. Computer Vision and Pattern Recognition(CVPR)*, pp.2599-2606, 2009.
- [6] Wei Zhang and Kosecka J., "Image Based Localization in Urban Environment," *3D Data Processing, Visualization, and Transmission, Third International Symposium(3DPVT)*, pp.33-40, 2006.
- [7] David P, "Detection of Building Facades in Urban Environments," *Proc. of the SPIE*, Vol. 6978, pp. 139-148, 2008.
- [8] Chang Chang Wu, "Repetition-based dense single-view reconstruction" *IEEE Int. Conf. Computer Vision and Pattern Recognition(CVPR)*, pp.3113-3120, 2011.
- [9] Jan Bohm, Nobert Haala, Peter Kapusy, "Automated Appearance-Based Building Detection in Terrestrial Images," *In ISPRS Commission V Symposium, International Archives on Photogrammetry and Remote Sensing, Vol. 34number 5*, pp. 491-495, 2002.
- [10] Yi Li, Shapiro. L.G, "Consistent line clusters for building recognition in CBIR," *Proc. of the 16th Int. Conf. Pattern Recognition*, Vol. 3, pp.952-956, 2002.
- [11] T. Deselaers, D. Keysers, H Ney, "Features for Image retrieval: an experimental comparison," *Image Retrieval, Vol. 11, No.2*, pp.77-107, 2008
- [12] N. Senthilkumaran, R. Rajesh, "Edge Detection Techniques for Image Segmentation – A Survey of Soft Computing Approach," in *International Journal of Recent Trends in Engineering*, Vol. 1, No. 2, May 2009
- [13] Tremereau. A, "Regions adjacency graph applied to color

- image segmentation,” *IEEE Trans. on. Image Processing*, Vol. 9, No. 4, pp. 735-744, 2000
- [14] Wang J.Z, Jia Li, “SIMPLIcity: Semantics-sensitive integrated matching for picture libraries,” *IEEE Trans. on. Pattern Analysis and Machine Intelligence*, Vol. 23, No. 9, pp. 947-963, 2001.
- [15] Y. Rubner, C. Tomasi, LJ Guibas, “The earth mover’s distance as a metric for image retrieval,” in *International Journal of Computer Vision*, Vol. 40, No. 2, pp. 99-121, 2000
- [16] Shao H, Svoboda T, Tuytelaars T, Gool L. “ZubuD – Zurich Buildings Database for Image Based Recognition,” *Computer Vision Lab, Swiss Federal Institute of Technology, Switzerland. Zurich*, 2003
- [17] Müller H, Müller W, Squire D M, Marchand-Maillet S, Pun T. “Performance Evaluation in Content-Based image Retrieval: overview and Proposals,” *Patter Recognition Letters*, Vol. 22, No. 5, pp. 593-601, April 2001.

이 민 철

1983 부산대학교 기계공학과
(공학사)

1988 쑤저우대학교 이공학 연
구과(공학석사)

1991 쑤저우대학교 물리공학
연구과(공학박사)

2000.8~2001.8 노스캐롤라이나주립대학교(NCSU) 방문
교수

2009.8~2010.8 퍼듀대학교 방문교수

1991~현재 부산대학교 기계공학부 교수

관심분야 : 시스템 규명, 로봇제어, 의료로봇, 지능형 서
비스로봇, 메카트로닉스

허 정 훈

2010 동의대학교 메카트로닉스
공학과(공학사)

2010~현재 부산대학교로봇협
동과정(석사과정)

관심분야 : 머신비전, 로봇
스, 자율주행 로봇