

OpenSlide 병리학 이미지에 Leveling 과 Tiling 기법을 적용한 유사 세포 이미지 탐색

이재구*, 고영웅**

*, **한림대학교 컴퓨터공학과

e-mail : yuko@hallym.ac.kr

Apply Leveling and Tiling to OpenSlide pathology image to retrieve similar cell image

Jae Gu Lee*, Youngwoong Ko**

*, **Dept. of Computer Engineering, Hallym University, Chuncheon, South Korea

요 약

다양한 분야에서 특정 이미지와 유사한 내용의 이미지를 찾기 위해서 이미지 유사도 기법을 지원하는 웹 또는 프로그램이 많이 사용되고 있다. 본 논문에서는 이미지 유사도 기법을 병리학 이미지에 적용하여 특정 세포를 찾는 데 사용함으로써 사용자에게 정확하고 신뢰성 있는 정보를 전달하려 한다. 실제로 병리학에서 특정 세포를 찾기에는 세포의 크기가 너무 작고 모양이 일정하지 않는다는 점 때문에 유사한 세포를 찾기가 힘들다. 이를 해결하기 위해 Leveling 과 Tiling 방식을 적용하여 데이터를 추출하고 이미지 동기화를 통해 유사도 비교 시 정확도를 높이도록 한다. 다양한 병리학 이미지 포맷을 지원하는 OpenSlide 기반의 대용량 이미지에 제안하는 방식을 적용하여 특정 세포와 유사한 부분을 찾아내는 실험을 통하여 제안하는 방식이 효율적임을 보인다.

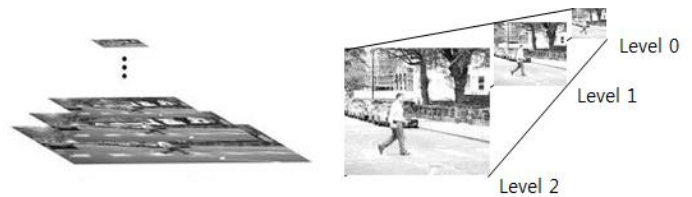
1. 서론

최근 이미지 분석 기술을 사용하는 분야가 늘어남에 따라 다양한 종류의 연구가 진행되고 있다. 이미지에서 특징을 추출하여 사람이라는 데이터와 유사한 특징을 가졌는지 확인하고 해당 이미지의 움직임 추적을 하는 기술 [1] 그리고 사람의 얼굴을 분석하는 기술 [2]이 연구되고 있다. 의학분야에서는 사람의 뇌를 CT 로 촬영하고 해당 이미지를 사용하여 데이터베이스에 저장되어 있는 유사한 뇌를 검출해내는 연구도 존재한다 [3]. 병리학 분야에서도 이미지 유사도 기술을 사용하려 하지만 여러 가지 문제점으로 인하여 유사도 기술을 적용하기란 쉽지 않다. 일반 분야와 달리 병리학 이미지에 유사도 기법을 적용하여 작은 암세포를 찾을 때는 정확도와 신뢰도가 완벽해야 한다. 그러나 이미지 유사도 기법은 항상 정확한 결과값을 유추하지 못한다 [4].

병리학 이미지는 지역 및 병원에 따라서 이미지의 포맷 형식도 다양하다. 유사한 연구에서는 다양한 병리학 이미지 포맷을 지원하는 OpenSlide [5]를 사용하여 이러한 문제를 해결한다. OpenSlide 는 하드웨어, 운영체제 그리고 웹 브라우저 종류에 상관없이 사용 가능하다. OpenSlide 는 9 가지의 slide 포맷을 지원하기 때문에 다양한 병리학 이미지를 다루기에도 최적이다.

그리고 여러 슬라이드 이미지를 압축하여 저장하기 때문에 이미지를 읽어 들일 때도 효율적이다.

또한 병리학 특성상 하나의 세포 이미지를 여러 번 촬영하기 때문에 다수의 이미지를 하나의 slide 이미지로 압축하는 그림 1 과 같은 방식을 사용하는 OpenSlide 가 병리학 이미지를 보관할 때도 매우 효율적이다.



(그림 1) 이미지 피라미드

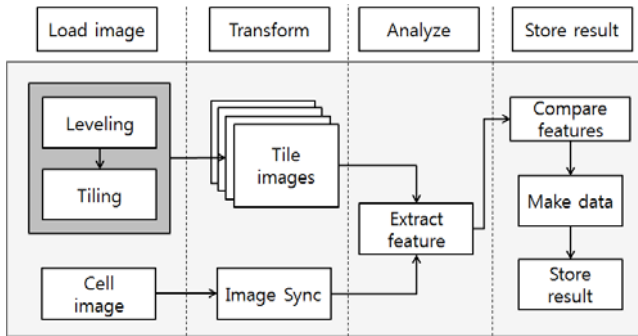
이미지 유사도를 측정할 때 특징 값을 구하는 방식은 분야와 사용되는 key 값 (색, 모양, 크기 등)이 무엇이나에 따라 여러 종류의 알고리즘이 존재한다 [6].

본 논문은 병리학 이미지에 유사도 비교 알고리즘을 적용했을 때의 문제를 해결하기 위하여 OpenSlide 기반의 대용량 이미지에 Leveling 과 Tiling 기법을 적용하여 정확하게 유사한 세포 이미지를 찾을 수 있도록 한다. 또한 이미지 크기 동기화를 통하여 이미지 유사도 작업 결과의 정확도와 신뢰도를 향상시킨다. 실험 결과를 통하여 제안하는 방법이 정확하게 세포

¹ This research was supported by Hallym University Research Fund 2015(H20150684)"

이미지를 찾아냄을 보인다.

2. 시스템 구성

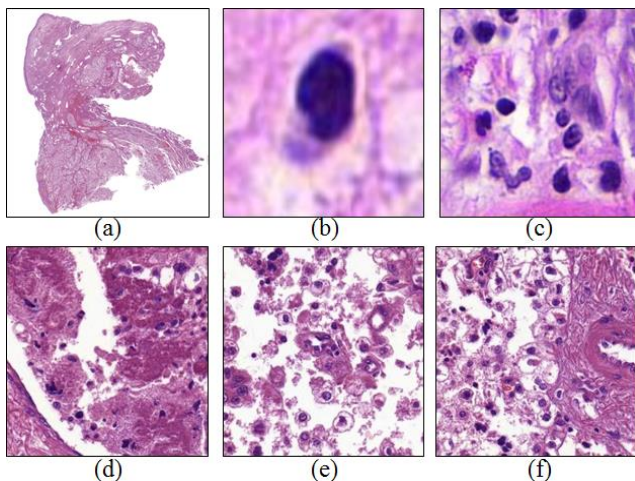


(그림 2) 시스템 순서도

그림 2는 제안하는 시스템의 순서도이다. OpenSlide 이미지가 촬영되면 Vips 패키지를 사용하여 Leveling과 Tiling을 적용하여 이미지들을 추출한다. Slide 이미지 중에서 비율이 유사한 Level의 이미지를 추출하는데 Leveling이 사용된다. 대용량 이미지일 경우 이미지 전체를 메모리에 등록할 수 없기 때문에 Tiling 기법을 적용하여 이미지의 크기를 고정적으로 분할한다. 비교 대상이 되는 세포 이미지는 Tiling된 이미지와 비교 작업을 수행했을 시 정확하고 빠른 비교작업을 수행하기 위해서 Image Sync 작업을 수행한다. Leveling과 Tiling이 적용된 이미지들은 Image Sync가 적용된 특정 세포 이미지와 비교될 때 마다 특징을 추출하여 비교된다. 비교 결과로 유사한 세포 이미지가 있다면 로그파일을 만들고 저장한다.

3. 실험 및 결과

본 실험은 Intel Core2 Duo CPU E7500 2.93GHz를 사용하는 개인용 컴퓨터에 Ubuntu 14.04 버전을 설치하고 GPU를 사용하지 않고 2GB의 ram을 장착하여 실험을 진행하였다. 실험에 사용된 영상은 OpenSlide Demo에서 사용되는 이미지를 받아서 참조한 2.4GB OpenSlide 이미지이다.



(그림 3) 유사 세포 탐색 결과

그림 3은 Image Sync 과정을 거친 특정 세포 이미지와 송신 받은 OpenSlide 이미지 중 level 3에 해당하는 이미지에 제안하는 방법을 적용하여 유사도를 비교한 결과이다. 그림 (a)는 12729 x 27497 크기의 level 3에 해당하는 OpenSlide 이미지이다. 추후 그림 (b)와 유사도 비교를 수행하여 그림 (a)에서 그림 (b)와 유사한 세포들을 찾아낸다. 원본 크기가 20 x 18인 특정 세포 이미지 (b)에 Image Sync를 수행하여 Tile 이미지의 사이즈 256 x 256과 동일한 크기로 변환한 후 Tiling된 수만 장의 Tile 이미지와 유사도 비교를 수행한다. 이미지 (c), (d), (e), (f)는 유사도 비교를 수행하여 찾은 수많은 유사한 이미지 중 일부를 출력한 결과이다. 결과를 보면 알 수 있듯이 Image Sync를 수행하여 픽셀 값이 늘어난 세포를 사용하여도 유사한 세포를 포함하는 Tile 이미지를 정확히 찾아내는 것을 볼 수 있다.

4. 결론 및 향후 연구

본 논문은, OpenSlide 기반의 대용량 병리학 이미지에 Leveling과 Tiling 기법을 적용하여 특정 Level의 이미지만 추출한 후 이미지 유사도 비교를 수행하기 쉽게 고정된 크기로 Tiling하는 이중 추출 방법을 사용한다. 또한 이미지 유사도 비교의 정확도를 향상시키기 위해 비교되는 실험 데이터 이미지의 픽셀 개수와 추출된 이미지의 크기를 동기화한다. 실험을 통하여 제안하는 방법에 단순한 이미지 유사도 알고리즘을 적용하였음에도 매우 정확하게 유사한 세포 이미지를 찾는 것을 보아 제안하는 방식이 매우 효율적임을 보인다. 추후 탐색된 결과 값이 원본 OpenSlide에 반영 되도록 이미지를 수정하는 작업과 metadata를 생성하여 관리자가 손쉽게 확인하고 수정할 수 있도록 추가 구현할 계획이다.

참고문헌

- [1] Corvee, Etienne, and Francois Bremond. "Body parts detection for people tracking using trees of histogram of oriented gradient descriptors." *Advanced Video and Signal Based Surveillance (AVSS)*, 2010 Seventh IEEE International Conference on. IEEE, 2010.
- [2] Déniz, Oscar, et al. "Face recognition using histograms of oriented gradients." *Pattern Recognition Letters* 32.12, 1598-1603, 2011.
- [3] Pan, Haiwei, et al. "Brain CT image similarity retrieval method based on uncertain location graph." *Biomedical and Health Informatics, IEEE Journal of* 18.2, 574-584, 2014.
- [4] Pan, Xunyu, and Siwei Lyu. "Region duplication detection using image feature matching." *Information Forensics and Security, IEEE Transactions on* 5.4, 857-867, 2010.
- [5] Goode, Adam, et al. "OpenSlide: A vendor-neutral software foundation for digital pathology." *Journal of pathology informatics*, 4.1, 27, 2013.
- [6] Yang, Hanxuan, et al. "Recent advances and trends in visual tracking: A review." *Neurocomputing*, 74.18, 3823-3831, 2011.