김일성종합대학학보
정보과학

주체109(2020)년
제66권 제3호

XY절단알고리듬과 중첩신경망에 의한 문서화상의 령역분할방법

리광철, 고진혁

문서화상으로부터 검색을 위한 정보들을 추출하고 분석하고 저장하기 위해서는 먼저 문서화상을 본문구역과 비본문구역에 해당한 령역들로 분할하여야 한다.

지금까지 문서화상의 령역분할을 위한 많은 방법[1, 2]들이 제안되였지만 대부분의 방법들은 문서화상의 내용이나 특징에 대한 사전지식을 필요로 하거나 문서화상의 질에 의존하는 결함이 있었다.

론문에서는 XY절단알고리듬과 중첩신경망을 리용하여 문서화상의 령역들을 류형별로 분할하기 위한 한가지 방법을 제안하였다.

1. XY절단알고리듬과 삽화검출에 의한 문서화상의 령역분할

문서화상에는 본문, 그림, 표 등 각이한 형식의 령역들이 존재한다. 그리하여 먼저 XY절단알고리듬[3]을 리용하여 문서화상에 대한 초기분할을 진행한다.

XY알고리듬은 반복적인 top-down알고리듬으로서 문서화상의 매 화소들을 수평 및 수직방향으로 투영하여 얻어지는 투영히스토그람에 기초하여 문서화상령역을 반복적으로 분할해나간다.

그러나 분할된 령역들에는 본문으로 된 령역뿐아니라 삽화와 본문이 함께 있는 령역들도 있다.

이로부터 다음의 알고리듬에 의하여 매 분할된 령역들에서 삽화령역을 검출한다.

본문령역과 삽화령역을 구별하기 위한 특징으로서 자체상관행렬을 리용한다.

문서화상을 크기가 $n\times n$인 블로크들로 분할하고 매 블로크에서 자체상관행렬을 다음의 식에 따라 계산한다.

$$C(k,\ l)=\sum_{y=\max(0,\ l)}^{n-1+\min(0,\ l)}\ \sum_{x=\max(0,\ k)}^{n-1+\min(0,\ k)} I(x,\ y)\cdot I(x+k,\ y+l) \qquad (1)$$

여기서 l과 k는 $[-n/2,\ n/2]$에서 정의되며 블로크의 크기 n은 본문블로크의 반복적인 패턴들이 강조되도록 설정한다.

식 (1)에 의한 자체상관계산은 많은 계산량을 필요로 하며 효률적이지 못하다.

자체상관함수의 푸리에변환은 에네르기스펙트르와 같고 결국 자체상관은 에네르기스펙트르의 역푸리에변환과 같다.

그러므로 블로크의 자체상관을 고속푸리에변환의 두단계 계산을 통하여 효과적으로 계산할수 있다.

블로크에 대한 자체상관은 무늬에서 의미가 있는 방향을 추정하는데 리용된다.

그러므로 $C(k,\ l)$을 극자리표계로 변환하고 다음의 식에 따라 방향히스토그람을 구한다.

$$w(\theta) = \sum_{r\in(0,\ n/2]} C(r\cos\theta,\ r\sin\theta) \tag{2}$$

자체상관행렬은 정의에 의하여 중심에 관하여 대칭이기때문에 $[0^o, 180^o)$ 범위내에서의 방향히스토그람만 고려한다.

매 블로크들을 방향히스토그람에 의하여 부호화하면 본문블로크들은 0^o와 180^o근방에서 봉우리가 나타나며 반대로 삽화블로크들은 다중양상형태로 나타난다.

최종적으로 방향히스토그람과 자체상관행렬을 수평 및 수직방향으로 투영하여 얻어지는 사영히스토그람을 련결하여 블로크에 대한 특징벡토르를 구성하고 SVM분류기를 리용하여 본문령역과 삽화령역을 판별한다.

검출된 삽화령역들을 제거하고 다시 XY절단알고리듬을 적용하여 최종적인 분할령역들을 얻는다.

2. 중첩신경망에 의한 문서화상의 령역류형분류

론문에서는 문서화상에 포함되는 령역의 류형을 크게 6가지 즉 본문, 그림, 그라프, 표, 수학식, 악보로 하였다.

령역류형의 분류는 중첩신경망을 리용하여 진행한다.

중첩신경망의 입력은 하나의 령역으로부터 추출한 크기가 $n\times n$인 블로크이며 출력은 우의 6가지 류형중 어느 하나의 류형이다.

령역의 류형분류를 위한 중첩신경망의 구성을 표에 보여주었다.

표. 령역의 류형분류를 위한 중첩신경망의 구성

층	핵크기	걸음크기	려파기수	입력	출력
conv1	11×11	1×1	32	(1, n, n)	(32, n, n)
mp1	2×2	2×2	−	(32, n, n)	(32, n/2, n/2)
conv2	5×5	1×1	32	(32, n/2, n/2)	(32, n/2, n/2)
mp2	2×2	2×2	−	(32, n/2, n/2)	(32, n/4, n/4)
conv3	3×3	1×1	32	(32, n/4, n/4)	(32, n/4, n/4)
mp3	2×2	2×2	−	(32, n/4, n/4)	(32, n/8, n/8)
fc1				(32 · n^2/64)	1 024
dropout				1 024	512
fc2				512	6

어떤 령역에 대한 분류를 하기 위하여 령역안의 매 블로크들을 학습된 중첩신경망에 입력하였을 때의 fc1층의 출력벡토르(1 024차원)를 얻고 령역안의 모든 블로크들에 대한 1 024차원의 특징벡토르들에 대한 평균과 공분산행렬을 각각 구한다.

령역의 기하학적자리표가 (R_x, R_y, R_w, R_h) 이고 블로크들의 모임 B를 포함하고있는 령역 R의 특징벡토르는 다음과 같다.

$$\boldsymbol{f}_R = \left[\frac{R_x}{W}, \frac{R_y}{H}, \frac{R_w}{W}, \frac{R_h}{H}, \frac{R_w}{R_h}, \mu(B), \sigma(B)\right] \tag{3}$$

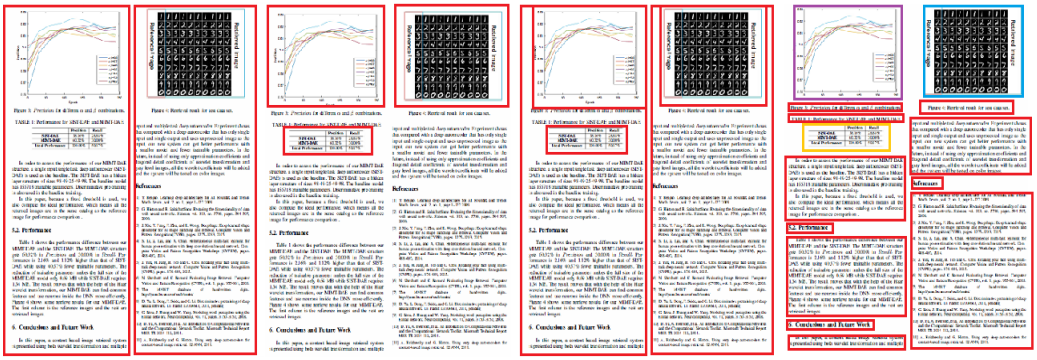
여기서 W, H는 각각 페지의 너비와 높이이며 $\mu(B), \sigma(B)$는 각각 B안의 모든 블로크들에 대한 평균벡토르와 공분산행렬의 왼쪽웃3각형원소들을 1차원으로 라렬한 벡토르이다.

령역에 대한 최종분류는 령역에 대한 특징벡토르 f_R와 우연숲분류기(Random Forest Classifier)[4]를 리용하여 진행한다.

3. 문서화상의 령역분할실험

실험에서 리용된 문서화상의 크기는 2 448×3 264이며 블로크의 크기는 64×64로 설정하였다.

문서화상에 대하여 XY절단알고리듬, 삽화검출알고리듬, 령역류형분류알고리듬을 각각 적용하였을 때의 결과를 그림에 보여주었다.



ㄱ) ㄴ) ㄷ) ㄹ)

그림. 문서화상의 령역분할실험결과

ㄱ) XY절단알고리듬적용결과, ㄴ) 삽화검출알고리듬적용결과,

ㄷ) XY절단알고리듬을 다시 적용한 결과, ㄹ) 령역류형분류결과

론문의 방법에 의한 문서화상의 령역분할정확도는 약 92%로서 XY절단방법이나 백색령역분석방법에 비하여 약 20% 더 높다.

맺 는 말

XY절단알고리듬과 중첩신경망을 리용하여 문서화상의 령역분할과 령역들의 류형결정을 동시에 진행하기 위한 한가지 방법을 제안하고 실험을 통하여 그 효과성을 검증하였다.

참 고 문 헌

[1] Pradipta Maji, Shaswati Roy; Applied Soft Computing, 1, 1, 2015.

[2] Y. L. Chen, B. F. Wu; Pattern Recognit, 42, 1419, 2009.

[3] Jaekyu Ha et al.; Proceedings of the Third International Conference on Document Analysis and Recognition, 2, 952, 1995.

[4] Archana Chaudhary et al.; Information Processing in Agriculture, 3, 4, 215, 2016.

주체109(2020)년 5월 5일 원고접수

Segmentation Method of Document Image by XY-cut Algorithm and CNN

Ri Kwang Chol, *Ko Jin Hyok*

In this paper, we propose a method for clustering regions of document image using XY-cut algorithm and CNN.

Keywords: XY-cut algorithm, CNN, document image, segmentation