

혼합가우스분포와 기대값최대화를 리용한 맹인글자 문서화상 행추출의 한가지 방법

최철, 이정철

위대한 령도자 김정일동지께서는 다음과 같이 교시하시였다.

《정보기술발전을 위한 과학연구사업을 앞세워야 합니다.》(《김정일선집》 증보판 제20권 380페이지)

맹인글자문서화상인식체계는 문서에 인쇄된 맹인용점글자(Braille Character)를 화상입력 장치로 읽어들이 인식하는 체계로서 최근시기 그 연구가 본격화되고있다.[2]

맹인글자문서화상의 인식에서는 행을 정확히 추출하는것이 뒤단계의 인식을 위해 매우 중요한 문제로 제기된다. 이러한 행추출방법에는 호우변환과 히스토그램을 리용한 방법, 라돈변환(RT: Radon Transform)에 의한 방법 등 여러가지가 있다. 논문에서는 혼합가우스분포를 리용한 맹인글자문서화상에서의 행추출의 한가지 방법을 제기하고 실험을 통하여 그 유효성을 검증하였다.

1. 혼합가우스분포를 리용한 행추출

논문에서는 부분구역선택에 의한 경사각추출방법[1]에 의해 문서의 경사보정은 이미 진행된것을 전제로 한다.

맹인글자문서화상을 인식하기 위해서는 맹인글자를 이루는 매개 행을 결정해야 하는데 이를 위해서는 문서의 웃가장자리로부터 첫행까지의 거리(b)와 행사이간격(β)을 추정해야 한다. 앞으로는 이 두 파라미터(b 와 β)를 행파라미터라고 표기한다.

맹인글문서는 그 특성상 하나의 행이 3개의 부분행으로 이루어지므로 행분포는 3개의 가우스분포의 혼합으로 모형화될수 있다. 그러나 우리는 모형의 복잡성과 파라미터추정에서의 느린 수렴성문제를 극복하기 위하여 한행안의 3개의 부분행들을 하나의 가우스분포로 모형화한다.

문서를 2진화하고 연결성분추출을 진행했을 때 매 연결성분의 수직자리표들의 분포는 다음과 같이 표현할수 있다.

$$g(y) = \frac{1}{L} \sum_{l=0}^{L-1} N(y; b + l\beta, \sigma) \quad (1)$$

식 (1)에서 L 은 문서안의 행개수, l 은 행번호, y 는 추출된 연결성분들의 수직자리표, b 와 β 는 행파라미터, σ 는 모형의 분산을 나타낸다.

식 (1)에서 보는바와 같이 모든 연결성분의 수직자리표 y 의 분포는 1차원가우스혼합분포로 표현할수 있다.

문제는 식 (1)이 최대가 되는 모형의 행파라미터를 추정하는것인데 우리는 이 문제를 기대값최대화(EM: Expectation Maximization)알고리즘으로 해결하려고 한다.

매 연결성분을 C_i 로 표현하면 기대값최대화문제는 다음과 같이 해결할수 있다.

1) 기대값결정단계

먼저 기대값최대화를 위한 우도함수를 다음과 같이 결정한다.

$$\begin{aligned} \log p(y_1, \pi_1, y_2, \pi_2, \dots, y_N, \pi_N) &= \log \prod_{i=1}^N p(y_i, \pi_i) = \\ &= \sum_{i=1}^N \log \left(\pi_i \left(c \frac{1}{L} \sum_{l=0}^{L-1} N(y_i; b+l\beta, \sigma) + (1-\pi_i) \frac{1}{S} \right) \right) \end{aligned} \quad (2)$$

식 (2)에서 N 은 연결성분개수, y_i 는 i 번째 연결성분의 수직자리표, c 는 계산상수, δ 는 모형의 분산, S 는 해당 연결성분이 잡음일 확률이다. 그리고 π_i 는 연결성분의 사전확률로서 다음과 같이 계산된다.

$$\begin{aligned} \pi_i &= p(C_i \in \text{dot} | F_i) = \\ &= \frac{p(F_i | C_i \in \text{점모임})p(C_i \in \text{점모임})}{p(F_i | C_i \in \text{점모임})p(C_i \in \text{dot}) + p(F_i | C_i \in \text{잡음모임})p(C_i \in \text{잡음모임})} \end{aligned} \quad (3)$$

한편 식 (2)를 그것의 파라미터들에 관하여 미분하고 그것들을 령으로 놓으면 복잡한 비선형방정식모임이 생기는데 이것을 풀기는 매우 어렵다. 그러므로 우리는 반복적인 EM방법을 리용하여 우도를 최대화하도록 파라미터를 결정한다. 이 경우에 우연변수들은 $\omega_1, \omega_2, \dots, \omega_N$ 이며 여기서 $\omega_i \in \{0, 1, \dots, L-1, L\}$ 은 행번호를 나타낸다.

기대값결정단계에서는 현재(t)시각에 추정 $\theta^t = \{b^t, \beta^t, \sigma^t\}$ 가 주어질 때 매 ω_i 의 확률분포를 구해야 한다. 즉

$$\begin{aligned} P_{il} &= P(\omega_i = l | y_i, \pi_i, \theta^t) = \frac{P(\omega_i = l, y_i | \pi_i, \theta^t)}{P(y_i | \pi_i, \theta^t)} = \\ &= \frac{P(\omega_i = l, y_i | \pi_i, \theta^t)}{\sum_{l=0}^L P(\omega_i = l, y_i | \pi_i, \theta^t)} = \frac{P(y_i | \omega_i = l, \pi_i, \theta^t)P(\omega_i = l | \pi_i, \theta^t)}{\sum_{l=0}^L P(y_i | \omega_i = l, \pi_i, \theta^t)P(\omega_i = l | \pi_i, \theta^t)} = \\ &= \frac{P(y_i | \omega_i = l, \theta^t)P(\omega_i = l | \pi_i, \theta^t)}{\sum_{l=0}^L P(y_i | \omega_i = l, \theta^t)P(\omega_i = l | \pi_i, \theta^t)}, \end{aligned} \quad (4)$$

$$P(\omega_i = l | \pi_i, \theta) = \begin{cases} \frac{1}{L} \pi_i & 0 \leq l \leq L-1 \\ 1 - \pi_i & l = L \end{cases}, \quad (5)$$

$$P(y_i | \omega_i = l, \theta) = \begin{cases} cN(y_i; b+l\beta, \sigma) & 0 \leq l \leq L-1 \\ \frac{1}{S} & l = L \end{cases}. \quad (6)$$

식 (4)에서 ω_i 가 주어질 때 y_i 와 π_i 의 조건부독립성을 리용한다. (ω_i 를 알고있다고 하면 y_i 는 $\omega_i < L$ 인 경우 유효한 점이고 $\omega_i = L$ 은 인위적인 점으로 된다.)

2) 최대화단계

우선 이 단계에서는 보조함수 $Q(\theta, \theta')$ 를 다음과 같이 만든다.

$$Q(\theta, \theta') = \sum_{i=1}^N \sum_{l=0}^{L-1} \log \left(-\log \sigma - \frac{1}{2} \frac{(y_i - b - l\beta)^2}{\sigma^2} \right) P_{il} + C \quad (7)$$

여기서 C 는 파라메터 $\theta = \{b, \beta, \sigma\}$ 에 대해 상수이며 따라서 최량화문제에 영향을 주지 않는다.

다음 보조함수의 파라메터들의 다음번 추정

$$\theta^{t+1} = \{b^{t+1}, \beta^{t+1}, \sigma^{t+1}\}$$

을 다음과 같이 결정한다.

$$\theta^{t+1} = \arg \max_{\theta} Q(\theta, \theta') \quad (8)$$

$Q(\theta, \theta')$ 를 최대화하기 위하여 파라메터 $\theta = \{b, \beta, \sigma\}$ 의 매개에 대해 Q 의 편도함수를 계산하여 령으로 되게 하여야 한다. 즉

$$\frac{\partial}{\partial b} Q(\theta, \theta') = 0 \Rightarrow \left(\sum_{i=1}^N \sum_{l=0}^{L-1} P_{il} \right) b + \left(\sum_{i=1}^N \sum_{l=0}^{L-1} l P_{il} \right) \beta = \left(\sum_{i=1}^N \sum_{l=0}^{L-1} P_{il} y_i \right), \quad (9)$$

$$\frac{\partial}{\partial \beta} Q(\theta, \theta') = 0 \Rightarrow \left(\sum_{i=1}^N \sum_{l=0}^{L-1} l P_{il} \right) b + \left(\sum_{i=1}^N \sum_{l=0}^{L-1} l^2 P_{il} \right) \beta = \left(\sum_{i=1}^N \sum_{l=0}^{L-1} l P_{il} y_i \right), \quad (10)$$

$$\frac{\partial}{\partial \sigma} Q(\theta, \theta') = 0 \Rightarrow \frac{\sum_{i=1}^N \sum_{l=0}^{L-1} P_{il} (y_i - b - l\beta)^2}{\sum_{i=1}^N \sum_{l=0}^{L-1} P_{il}}. \quad (11)$$

방정식 (9)와 (10)은 다음과 같이 결합된다.

$$A \begin{bmatrix} b \\ \beta \end{bmatrix} = B, \quad \begin{bmatrix} b \\ \beta \end{bmatrix} = A^{-1} B \quad (12)$$

이때 행렬 A 와 B 는 다음과 같이 정의된다.

$$A = \begin{bmatrix} \sum_{i=1}^N \sum_{l=0}^{L-1} P_{il} & \sum_{i=1}^N \sum_{l=0}^{L-1} l P_{il} \\ \sum_{i=1}^N \sum_{l=0}^{L-1} l P_{il} & \sum_{i=1}^N \sum_{l=0}^{L-1} l^2 P_{il} \end{bmatrix}, \quad B = \begin{bmatrix} \sum_{i=1}^N \sum_{l=0}^{L-1} P_{il} y_i & \sum_{i=1}^N \sum_{l=0}^{L-1} l P_{il} y_i \end{bmatrix}^T \quad (13)$$

위의 식들을 리용하여 문서화상의 행파라메터들 $\theta = \{b, \beta, \sigma\}$ 가 추정되면 행의 위치를 정확히 알아낼수 있으며 이때 중간선위치는 다음과 같이 얻을수 있다.

$$y_l = b + l\beta - \frac{\beta}{2}, \quad (l=0, 1, \dots, L-1) \quad (14)$$

실천에서는 행개수를 1~30사이로 놓고 매개 행에서 최대우도를 내는 행파라메터를 계산한 다음 가장 큰 우도를 가질 때의 행개수와 행파라메터를 최종결과로 한다.

2. 실험결과 및 성능평가

제안된 맹인글자문서화상의 행추출방법의 효과성을 검증하기 위하여 분해능이 200dpi 인 화상입력장치(《BENQ-Scanner-5005》)로 입력한 50개의 문서화상을 $-20, -15, -10, -5, +5, +10, +15, +20^{\circ}$ 로 회전시켜 모두 400개의 실험화상을 준비하였다.

여기에 기초하여 실험을 진행하고 제안한 방법과 선행한 방법을 비교한 결과는 표와 같다.

표. 행추출률 평가

적용방법	문서개수	정확히 추출한 개수	행추출률/%
선행한 방법[2]	400	392	98.4
제안한 방법	400	398	99.5

표에서 보여준것처럼 제안한 방법의 행추출성능이 선행한 방법보다 높은것은 선행한 방법과는 달리 문서의 경사각을 추정파라미터로 결정하지 않았기때문이다.

실험에서는 행파라미터의 초기값을 $b=10, \beta=(\text{화상높이})/L, \sigma=\beta/4$ 로 주었다.

그림 1에 론문에서 제안한 방법을 리용한 행추출의 한가지 실패를, 그림 2에 행개수에 따르는 우도변화를 보여주었다.

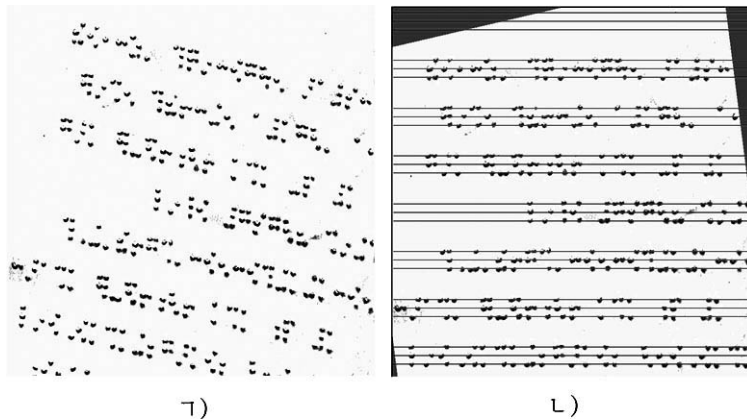


그림 1. 경사보정 및 행추출결과

ㄱ) 원화상, ㄴ) 추출결과

그림 2는 문서안에 행이 모두 25개까지 있을수 있다고 가정했을 때 행개수에 따르는 우도값을 표시한 그래프이다.

그림에서 행개수가 13일 때 우도값이 최대로 되므로 이때의 행파라미터를 리용하여 행추출을 진행하였다.

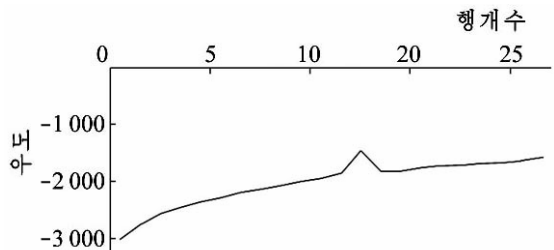


그림 2. 행개수에 따르는 우도변화

참 고 문 헌

[1] 김일성종합대학학보(자연과학), 52, 8, 40, 주체95(2006).

[2] M. Yousefi et al.; A Robust Probabilistic Braille Recognition System, Springer, 14, 2011.

주체104(2015)년 12월 5일 원고접수

**A Method for Extracting Lines of the Braille Document
Image using GMM and EM**

Choe Chol, Ri Jong Chol

We proposed a method to exactly extract the lines of braille document image using GMM and EM.

By experiments, we confirmed that the proposed method was very effective for extracting lines in the braille document images.

Key words: braille, GMM