

얼굴화상의 기미검출에서 지지벡토르기계의 특징을 합리적으로 선택하기 위한 한가지 방법

김광천, 리명철

경애하는 최고령도자 김정은동지께서는 다음과 같이 말씀하시였다.

《우리는 과학기술을 발전시키는데서 남들이 걸은 길을 따라만 갈것이 아니라 단계를 뛰어 넘어 비약적인 발전을 이룩하여야 합니다.》

론문에서는 얼굴화상에서 기미를 검출할 때 오류률을 증가시키는 머리칼영역에 대한 판정오차를 줄이기 위한 방법을 연구하였다. 우리는 기미영역을 판정하는데 지지벡토르기계(SVM)를 분류기로 리용하였으며 분류기의 판별성능을 높이기 위하여 특징을 합리적으로 선택하는 한가지 방안을 제안하였다.

선행연구[1]에서는 기미를 포함하는 피부영역과 기미가 없는 피부영역을 판단하는데 SVM분류기를 리용하였으며 특징으로는 회색준위동시발생행렬(GLCM)에 기초한 피부질평가를 위한 통계량들을 리용하였으나 머리칼이 없는 얼굴의 특정영역에서의 기미검출에 대하여 논의하고 머리칼영역이 기미검출에 미치는 영향에 대하여 논의하지 못하였다.

선행연구[2]에서는 려과기출력화상에 대하여 턱값판정방법을 리용하여 2진화상처리를 진행하고 련결성분의 크기에 기초하여 기미여부판단을 한다. 그러나 이 방법은 화상의 미세한 변화에도 큰 영향을 받으며 특히 머리칼영역이 기미로 잘못 검출되는 것과 같은 오류가 있다. 여기서도 얼굴영역에 머리칼이 없는 얼굴화상자료기지에 대하여 성능평가를 진행하고있다.

우리는 머리칼이 얼굴영역의 일부를 가리우거나 드리우는데 대한 제한을 주지 않은 상태에서 기미영역을 검출하는 방법을 연구하였다. 이때의 문제점은 머리칼부분이 기미영역으로 잘못 판단되는것이다.

기미판정에서의 오류는 기미영역을 기미가 없는 영역으로 판단하는 1종의 오류와 기미가 없는 영역을 기미영역으로 판단하는 2종의 오류가 있다.

우리는 머리칼이 있는 얼굴영역에 대하여 두가지 종류의 오류를 모두 감소시키면서도 2종의 오류를 감소시키는데 더 큰 비중을 두고 성능을 높이기 위한 방법을 연구하였다.

우리는 선행연구[2]의 방법과 류사한 방법으로 기미후보를 얻지만 턱값판정조건을 약화시켜 상대적으로 많은 후보를 얻은 다음 SVM분류기로 후보가 기미영역인가를 판정하는 방법을 제안하였으며 이때 기미없는 머리칼영역이 기미영역으로 판정되는 오류를 줄일수 있는 특징선택방법을 제안하고 실험결과를 주었다.

우리는 SVM분류기를 학습시키기 위한 자료를 머리칼과 기미가 있는 영역 G_1 , 머리칼이 있고 기미가 없는 영역 G_2 , 머리칼이 없고 기미가 있는 영역 G_3 , 머리칼과 기미가 없는 영역 G_4 와 같이 4가지로 세분하였다.

이 4가지 표본들중에서 G_1 , G_3 과 G_2 , G_4 를 판별하는 분류기를 만들어야 한다.

특징은 선행연구[1]에서와 같이 피부특성을 나타내는 통계적특징을 밝기, Cr, angleA의 색공간에서 구한 42개 특징중에서 기미가 없는 머리칼영역이 기미로 판단되는 오류를 줄이는데 좋은 특징들만을 선택하여 리용하였다. 특징은 G_2 와 G_3 에 대하여 무리안에서는 변동이 작고 무리사이에는 변동이 큰 특성을 가지는것을 우선적으로 선택하는 방법으로 정하였다.

i 번째 특징의 무리 X 와 무리 Y 에 대한 F -특점[3]은

$$F(i)^{(X, Y)} = \frac{(\bar{x}_i^{(X)} - \bar{x}_i^{(X \cup Y)})^2 + (\bar{x}_i^{(Y)} - \bar{x}_i^{(X \cup Y)})^2}{(\sigma_i^{(X)})^2 + (\sigma_i^{(Y)})^2}$$

과 같이 계산한다. 여기서 $\bar{x}_i^{(X)}$ 와 $\sigma_i^{(X)}$ 는 각각 무리 X 에서 i 번째 특징의 평균과 표준편차이고 $X \cup Y$ 는 무리 X 와 Y 의 합이다.

실례로 $|G_k| = n_k$ ($k=1, \dots, 4$) 라고 놓았을 때 무리 G_1 의 평균과 표준편차, 무리 $G_1 \cup G_2$ 의 평균은 각각 다음과 같이 계산된다.

$$\bar{x}_i^{(G_1)} = \frac{1}{n_1} \left(\sum_{j=1}^{n_1} x_{ij}^{(G_1)} \right), \quad \sigma_i^{(G_1)} = \sqrt{\frac{1}{n_1} \left(\sum_{j=1}^{n_1} (x_{ij}^{(G_1)} - \bar{x}_i^{(G_1)})^2 \right)}, \quad \bar{x}_i^{(G_1 \cup G_2)} = \frac{1}{n_1 + n_2} \left(\sum_{j=1}^{n_1} x_{ij}^{(G_1)} + \sum_{j=1}^{n_2} x_{ij}^{(G_2)} \right)$$

우리는 특징을 합리적으로 선택하기 위하여 표본들을 4가지로 세분한 기초우에서 득점을 계산하였다.

오류률을 분석해보면 G_2 에 속하는 영역을 기미로 잘못 분류하는것이 많은 비중을 차지하므로 G_2 와 기미를 나타내는 G_1 , G_3 사이의 F -특점을 리용하여 새로운 득점을 $FF(i) = w_1 F(i)^{(G_2, G_1)} + w_2 F(i)^{(G_2, G_3)}$ 과 같이 계산한다. 여기서 w_1 과 w_2 는 G_2 , G_1 에 대한 득점과 G_2 , G_3 에 대한 득점에 각각 어느 정도의 무게를 주겠는가를 나타내는 무게결수이다.

문헌에서는 G_2 와 G_3 을 잘 분리하도록 하는데 보다 큰 의의를 부여하고 최적의 w_1 과 w_2 를 찾을 때 w_2 를 w_1 보다 크게 설정하였다. 이 득점에 기초하여 42개의 특징으로부터 n 개의 특징을 선택하는 방법은 다음과 같다.

① 검사자료에서 계산된 특징으로부터 득점계산을 진행하고 감소순서를 정한다.

② 순서화된 특징들에 대하여 앞에서부터 몇개를 취한것들중에서 분류성능이 가장 높은것을 선택한다. 다시말하면 검사자료에서 특징 f_i 로부터 득점 $FF(i)$ 를 계산하고 특징들을 득점의 감소순서로 순서화한것을 g_1, g_2, \dots, g_{42} 라고 하며 특징 g_1, g_2, \dots, g_k 에 대하여 4점교차확인을 진행하고 정확도를 a_k 라고 한다. $n = \arg \min_k a_k$ 라고 하면 최종특징

으로 g_1, g_2, \dots, g_n 을 선택한다.

이때 매 특징의 득점은 계산된 후 크기비교에만 참가하므로 전체를 w_1 로 나누고 $w = w_2 / w_1$ 를 리용하면 특징득점은 w 만을 가지고 $FF(i) = F(i)^{(G_2, G_1)} + w F(i)^{(G_2, G_3)}$ 과 같이 계산하면 된다.

우의 방법으로부터 특정된 w 에 대한 최적의 특징모임을 얻으며 w 를 일정한 범위에서 변화시키면서 오류가 최소로 되는 무게결수를 얻는다.

세분을 진행하지 않고 기미와 기미가 아닌 무리에 대한 F -특점[3]을 리용할 때와 비교하자.

분류의 정확도는 $ACC = (N_{f13} + N_{t24}) / (N_{13} + N_{24})$, 1종의 오류는 $ERR_A = N_{f13} / N_{13}$ 으로 계산되며 2종의 오류는 $ERR_B = N_{f24} / N_{24}$ 로 계산된다. 여기서 N_{13} 은 기미표본의 개수로서 G_1, G_3 에 속하는 표본의 개수 즉 $n_1 + n_3$ 과 같다. N_{f13} 은 실지로 기미인데 기미가 아니라고 판별된 표본의 개수이고 N_{t13} 은 실지로 기미인데 기미로 정확히 판별한 개수이다. 그리고 N_{24} 는 기미가 아닌 표본의 개수로서 $n_2 + n_4$ 와 같다. N_{f24} 은 실지로 기미가 아닌데 기미로 판별된 표본의 개수이고 N_{t24} 은 기미가 아닌것으로 정확히 판별된 개수이다.

그런데 논문에서는 기미가 아닌 머리칼령역이 기미령역으로 판별되는데 주목하고있으므로 2종의 오류는 $ERR_B_alt = N_{f2} / N_2$ 로 평가한다. 여기서 $N_2 = G_2 \models n_2$ 이고 N_{f2} 는 실지로 기미가 아닌 머리칼령역인데 기미로 판별되는 개수이다.

실험결과를 비교하면 표와 같다.

표. 비교결과			
	ACC	ERR_A	ERR_B_alt
제안한 방법	0.839	0.242	0.204
선행방법[3]	0.775	0.251	0.330

표에서 보는바와 같이 논문에서 제안한 방법은 선행연구[3]의 방법보다 분류의 정확도가 더 높고 2종의 오류가 줄어든다는것을 알수 있다.

참 고 문 헌

- [1] C. Y. Chang et al.; Dermatological Sciences and Applications, 3, 28, 2013.
- [2] C. C. Hsieh et al.; Journal of Computers, 10, 1, 12, 2015.
- [3] C. Y. Chang et al.; Pattern Recognition, 43, 10, 3494, 2010.

주체108(2019)년 6월 10일 원고접수

A Method to Select Features of Support-Vector-Machine Reasonably for Spot Detection from Facial Images

Kim Kwang Chon, Ri Myong Chol

We select features from gray level co-occurrence matrix reasonably to get higher classification accuracy when we detect spots from facial images including hair by Support-Vector-Machine.

Key words: spot detection, feature selection