Good Bank Evaluation by Chernoff Face Analysis using SAS macro faces

Jeongeun Lee a · Hyeseon Jeong a · Minji Kim a · Jihyun Kim a · Young Sook Son a,1

^aDepartment of Statistics, Chonnam National University

(Received September 27, 2013; Revised November 09, 2013; Accepted November 25, 2013)

Abstract

The SAS macro faces program by Friendly (1992) is for Chernoff face analysis, which is one of methods for the visualization representation of multivariate data. In this paper, we examined 18 face features used in the program and presented the modified program depending on the definition of a good face in days present. In addition, a good bank evaluation for 15 domestic banks was performed through Chernoff face analysis based on 11 bank economic indicators representing stability, the consumer satisfaction, soundness, and banks profitability.

Keywords: Chernoff face, SAS macro faces, good bank evaluation.

1. 서론

다변량 자료의 시각화는 자료들 간의 상호 종속관계, 특징적인 패턴, 군집, 이상치 등을 탐지하는 데 도움이 되므로 다양한 분야의 자료 분석에서 중요한 역할을 하고 있다. 다변량 자료를 2차원 평면상에 표현하는 다변량 시각화 표현방법으로는 glyphs (Anderson, 1960), profiles (Bertin, 1967), stars plot (Goldwyn 등, 1971), Andrews plot (Andrews, 1972), faces (Chernoff, 1973), MDS (Multidimensional Scaling; Kruskal과 Wish, 1978) 등이 있다.

체르노프 얼굴(Chernoff face)은 다변량 시각화 표현 방법들 중 하나로서 Chernoff (1973)가 다차원 자료를 이차원 평면상에 표현하기 위해 처음으로 인간의 얼굴을 사용하여 알려졌다. 체르노프 얼굴은 다변량 자료의 각 변수를 인간 얼굴의 눈, 코, 입 등의 18개 얼굴 특징에 대응시켜 각 관측치를 각기 다른 얼굴로 표현한다. Chernoff (1973)에서는 자메이카의 에오세 석회암 지층으로부터 고생시대의 87개 원생동물 화석을 관측하여 배아실에 관련된 6개의 변수에 얼굴 특징을 대응시켜 87개 화석의 얼굴을 이미지에 따라 군집화를 시도하였다. 또 다른 예로서, 미국 콜로라도 산의 산허리에서 4500foot 파 들어간핵심부에서 등 간격으로 채취한 53개 토양 표본에서 12개 광물 성분 값을 측정하였다. 12개 광물성분에 얼굴의 특징을 대응시켜 53개 토양 표본의 얼굴을 이미지에 따라 군집화 하였다.

오늘날까지 유명하게 회자되고 있는 체르노프 얼굴의 예는 Eugene Turner (1979)가 만들어 『Life in Los Angeles 1970』으로 제목이 붙은 지도이다. 이것은 미국 로스앤젤레스 지역의 삶의 질을 나타낸 지

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(NRF-2011-0022864).

¹Corresponding author: Professor, Department of Statistics, Chonnam National University, 300 Yongbong-Dong, Buk-Gu, Kwangju, 500-757, Korea. E-mail: ysson@jnu.ac.kr



Figure 1.1. The curvature of mouth and the direction of eyebrows

도로서 부의 정도, 실업률, 스트레스 지수, 백인의 비율과 같은 4개 변수를 얼굴 모양, 입의 곡률(곡선의 방향), 눈썹 기울기, 그리고 얼굴 색깔과 같은 얼굴 특징에 대응시켜 로스앤젤레스의 각 소지역을 체르 노프 얼굴로 표현하였다. 이 지도에서 행복한 얼굴을 보여주는 지역은 삶의 질이 좋은 지역으로 해석되 었다.

Huff 등 (1981)은 기업의 재정 상태를 파악할 수 있는 총자산 이익률, 차입금 상환, 현금 유통, 총자본, 유동성 비율 등 11개 재정상태 변수에 얼굴의 특징을 대응시켜 5개 기간에 걸친 2개 기업의 재정 상태를 체르노프 얼굴로 표현하였다. 결국 마지막에 파산한 회사는 파산하기까지 5개 기간에 걸쳐 점점 더 찡그리고 슬픈 얼굴로 변화해 갔으며, 다른 건실한 회사는 행복한 얼굴로 변화하는 모습을 보여주었다.

Spinelli와 Zhou (2004)는 체르노프 얼굴 분석을 통하여 미국인의 삶의 질을 나타내는 지도를 완성하였다. 삶의 질을 결정한다고 보는 6개 변수인 이혼율, 여성노동율, 대학 학위, 범죄율, 실업률, 평균 가구수입에 얼굴의 특징으로서 코의 크기, 눈썹의 방향, 얼굴 색, 귀의 크기, 입의 곡률, 그리고 머리카락의양을 각각 대응시켰다.

트리할로메탄(Trihalomethan; THM)류는 유기물을 함유한 원수(original water)를 염소 소독하는 과정에서 발생한다. 폴란드에서는 음용수에 녹아있는 THM 농도를 오랫동안 관측해 왔다. Astell 등 (2006)은 1993년~2000년 기간 동안 폴란드 12개 광역 구역에서 채취된 음용수에서 관측한 1,756개의 오르가노할로겐 화합물 중 주요한 5개 THM류 화합물에 입의 곡률, 코의 길이, 코의 넓이, 눈썹의 방향, 눈썹의 길이와 같은 얼굴의 특징을 대응시켜 12개 광역 구역을 체르노프 얼굴로 표현하였다. 각구역의 얼굴을 비교하여 원수 공급원(지상수 혹은 지하수)에 따라 THM 농도가 비슷한 구역별로 2개 그룹으로 나눌 수 있었다. 즉, 체르노프 얼굴분석은 THM 농도와 음용수 공급원의 지질학적 위치의 관계를 밝혀주었다.

체르노프 얼굴분석을 적용한 또 다른 예제로서 실험실 자료에서 변화추적 (Lott와 Durbridge, 1990), 브랜드에 대한 소비자 인식 (Golden과 Sirdesai, 1992), 환경보호정책에 대한 태도 (Apaiwongse, 1995) 등이 있다.

이상과 같은 체르노프 얼굴 분석의 예에서는 다변량 자료를 체르노프 얼굴을 이용해 시각적으로 표현함으로써 직관적으로 쉽고 빠르게 각 개체들의 특성을 파악할 수 있으며, 비슷한 얼굴을 가지는 개체들 끼리 군집화 하여 각 군집의 특징을 비교할 수 있다. 또한 체르노프 얼굴 분석은 glyphs, profiles, stars plot, Andrews plot과는 달리 행복한 얼굴 혹은 슬픈 얼굴처럼 인간의 정서적 감정을 유발하는 정보까지 준다. 그러나 바로 이 점 때문에 체르노프 얼굴분석을 하는 데 있어서 어려움이 존재한다. 즉, 얼굴의 희로애락에 가장 영향을 주는 얼굴 특징은 입의 곡률과 눈썹의 방향이므로 이 두 가지 얼굴 특징에 대응되는 변수에 따라 얼굴 이미지가 좌우된다. 체르노프 얼굴분석에서 변수가 순서 자료(ordinal data)인 경우는 순서를 반영하여 얼굴의 정서를 표현해야 한다. 예를 들면 실업률이 낮은 수준, 높은 수준에 따라 입의 곡률을 Figure 1.1의 (a), (b)로 각각 대응시키거나 혹은 눈썹의 방향을 Figure 1.1의(c), (d)와 같이 각각 대응시킨다면 변수의 의미상 순서와 얼굴의 정서가 일치한다. 그러나 Astell 등 (2006)에서는 암을 유발하는 화합물인 클로르포름이 고농도일수록 입의 곡률이 행복한 얼굴이 되도록

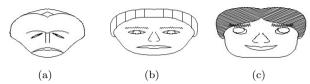


Figure 2.1. Chernoff faces: (a) all minimum (b) all average (c) all maximum

대응시켜 수질이 좋은 지역으로 착각이 있을 수 있다. 미국인의 삶의 질을 평가하기 위한 Spinelli와 Zhou (2004)의 체르노프 얼굴분석에서는 일하는 여성의 비율이 51%-59%, 64%-69%인 경우에 각각 Figure 1.1의 (c), (d)로 대응시켰다. 이 경우 여성 노동력이 높아질수록 화난 얼굴 이미지를 갖는다는 것은 설득력이 떨어진다. 즉, 체르노프 얼굴 분석에서는 데이터의 의미가 얼굴의 정서와 무관한 경우에도 얼굴의 정서로 표현될 수 밖에 없다.

Chernoff (1973) 이래 40여년이 흐른 지금까지 체르노프 얼굴분석에 관한 연구는 상대적으로 매우 적다. 이것은 위에서 설명한 바와 같이 주관적 변수 선택에 따라 표현되는 얼굴의 감정표현이 민감하기 때문이라 추측된다. 그럼에도 불구하고 체르노프 얼굴분석의 예제를 살펴보면 체르노프 얼굴분석은 전문적인 지식이 없어도 직관적으로 이해하기 쉬워 다양한 분야에서 널리 사용될 수 있는 흥미로운 통계적기법이라는 것을 알 수 있다.

통계패키지 SAS를 사용한 체르노프 얼굴 분석은 SAS macro 프로그램인 faces에 의해서 구현될 수 있다. 본 연구에서는 Friendly (1992)가 만든 SAS macro faces(Version 1.6-0 updated on 28 Sep 2011) 프로그램에서 사용된 18개 얼굴 특징변수를 살펴보고 실제 사용하면서 발견한 문제점에 기초하여 수정된 프로그램을 제안하였다. 또한 수정된 SAS macro faces 프로그램을 사용하여 은행의 안정성, 소비자 만족도, 건전성, 수익성을 나타내는 11개 은행경제지표에 기초한 체르노프 얼굴분석을 통하여 국내 15개 은행에 대하여 소비자 입장에서 좋은 은행 평가를 수행하였다.

2. SAS macro faces 프로그램의 수정

SAS macro 프로그램은 SAS를 사용자 정의형식으로 확장하기 위한 프로그램으로서 동일한 통계 작업을 시행하기 위해 반복해야 하는 프로그램의 분량을 줄일 수 있다. Friendly (1992)의 SAS macro faces 프로그램은 전체 742줄로 이루어져 있다. faces를 이용해서 체르노프 얼굴을 그릴 때 사용되는 18개 얼굴 특징변수들은 눈의 크기, 눈동자의 크기, 눈동자의 위치, 눈의 경사, 눈의 수평위치, 눈의 수직위치, 눈썹의 곡률, 눈썹의 밀도, 눈썹의 수평위치, 눈썹의 수직위치, 윗머리의 모양, 아랫머리의 모양, 얼굴의모양, 머리숱의 진한 정도, 머리카락의 방향, 코의 모양, 입의 크기, 입의 곡률 등이다. 이때 좌우의 얼굴 특징에 서로 다른 변수를 대응시킬 수 있으므로 최대 36개의 변수를 대응시킬 수 있다. 그러나 좌우에 서로 다른 변수를 대응시키는 경우 관측치의 군집화가 매우 복잡해질 뿐만 아니라 기형적인 얼굴 형상으로 표현되어 실제로 현실적이지 못하다. 따라서 본 연구에서는 좌우 대칭인 체르노프 얼굴 분석을 대상으로 한다.

체르노프 얼굴은 0에서 1까지의 실수 값을 가지는 얼굴 특징들이 모여 서로 다른 얼굴이 만들어진다. Figure 2.1은 체르노프 얼굴 이미지의 차이를 확인하기 위해서 18개의 모든 변수에 (a) 최솟값 0, (b) 평균값 0.5, (c) 최댓값 1을 각각 입력한 결과이다. 최솟값을 가진 (a)의 체르노프 얼굴은 눈, 코, 입의 크기가 작고 불만스러운 인상을 보여준다. 반면에 최댓값 1의 값을 가진 (c)의 얼굴은 눈, 코, 입의 크기가 클 뿐 아니라 여러 변수들이 조화를 이루어 밝고 자신감이 있어 보인다.

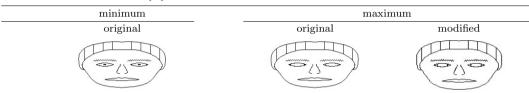
Table 2.1부터 Table 2.10에는 Friendly (1992)가 정의한 18개 얼굴 특징변수의 original image가 표

Table 2.1. Variable 1: size of eyes

| minimum | max | maximum | |
|----------|----------|----------|--|
| original | original | modified | |
| | | | |

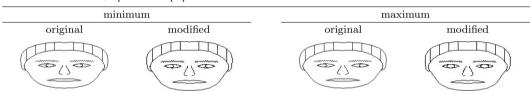
minimum: the size of eyes is small. maximum: the size of eyes is large.

Table 2.2. Variable 2: size of pupils



minimum: the size of pupils is small. maximum: the size of pupils is large.

Table 2.3. Variable 3: position of pupils



minimum: the position of pupils is far. maximum: the position of pupils is proper.

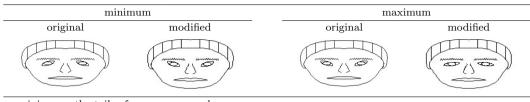
현되어 있다. 각 변수의 특징을 효과적으로 비교하기 위해서 해당하는 변수에 최솟값(0) 혹은 최댓 $\mathrm{ac}(1)$ 을 대입하고 해당하는 변수 이외의 값은 0.5를 대입하여 다른 변수의 영향을 배제시켰다. 또한 수정의 필요가 있는 얼굴 특징변수의 경우는 modified image를 추가하였다.

Table 2.1의 변수 1(눈의 크기)에서 눈동자의 크기에 맞지 않게 눈의 크기가 너무 큰 경우 좋은 인상이라고 볼 수 없기에 최댓값(1)의 눈의 크기가 0.8 정도 크기의 눈이 되도록 수정하여 눈동자의 크기에 적정한 눈의 크기가 될 수 있도록 하였다. SAS macro faces 프로그램의 687번째 줄에서 $p\{1\} = p\{1\} * 0.8를 추가하였다.$

Table 2.2의 변수 2(눈동자의 크기)에서 눈의 크기에 맞지 않게 눈동자의 크기가 너무 커서 흰자위만 있는 것처럼 보이는 경우에 좋은 인상이라고 볼 수 없다. 따라서 최댓값(1)의 눈동자 크기가 0.75 정도 크기의 눈동자가 되도록 수정하여 눈의 크기에 적정한 눈동자의 크기가 될 수 있도록 수정하였다. SAS macro faces 프로그램의 724번째 줄에서 $p\{2\} = p\{2\} * 0.75$ 를 추가하였다.

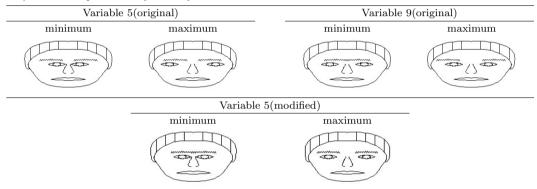
Table 2.3의 변수 3(눈동자의 위치)에서 최솟값(0)의 눈동자가 코에 가깝게 몰리고, 최댓값(1)의 눈동자가 서로 반대방향으로 몰려서 마치 사시처럼 얼굴이 표현된다. 하지만 사시를 갖는 얼굴이 좋은 인상이라고 볼 수 없다. 따라서 최솟값(0)의 얼굴이 사시처럼 표현되도록 하고 최댓값(1)의 얼굴에서 눈동

Table 2.4. Variable 4: slant of eyes



minimum: the tails of eyes are upward.
maximum: the tails of eyes are downward.

Table 2.5. Variable 5: horizontal position of eyes, Variable 9: horizontal position of eyebrows, Variable 5(modified): horizontal position of eyes and eyebrows



minimum: the horizontal position of eyes and eyebrows is close. maximum: the horizontal position of eyes and eyebrows is proper.

자의 위치가 눈의 중앙(0.5)에 위치하도록 수정하였다. SAS macro faces 프로그램의 727번째 줄에서 $p{3} = 1 - p{3} * 0.5$ 를 추가하였다.

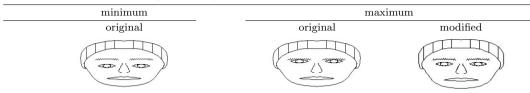
Table 2.4의 변수 $4(\pm 9)$ 경사)에서 최솟값(0)의 눈꼬리가 아래쪽으로 쳐져서 순한 인상으로, 최댓 값(1)의 눈 꼬리가 위를 향해서 사나운 얼굴로 보인다. 따라서 최솟값(0)의 경우 눈 꼬리가 위를 향하도록, 최댓값(1)의 경우 눈 꼬리가 아래를 향하되 적당히 아래로 향한 0.3의 값이 가장 좋은 얼굴이 되도록 하였다. SAS macro faces 프로그램의 682번째 줄에서 $p\{4\}=1-p\{4\}*0.7$ 를 추가하였다.

Table 2.5에서 변수 5(눈의 수평위치)와 변수 9(눈썹의 수평위치)의 original image를 보면 최댓값(1)에서 눈혹은 눈썹의 수평위치가 눈썹 혹은 눈의 위치에 관계없이 최대한 반대 방향을 향하고 있어 균형적인 얼굴이라고 볼 수 없다. 따라서 변수 5와 변수 9를 통합하여 눈과 눈썹의 수평위치가 같이 움직일 수 있도록 하였고, 눈이 너무 벌어진 얼굴이 좋은 인상이 아니기 때문에 벌어진 정도를 1에서 0.7로 수정하여 덜 벌어지도록 하였다. SAS macro faces 프로그램에서 $p\{9\}$ 를 모두 $p\{5\}$ 로 수정한 후, 678번째 줄에서 $p\{5\} = p\{5\} * 0.7$ 를 추가하였다.

Table 2.6에서 변수 6(눈의 수직위치)의 값이 커질수록 눈과 눈썹이 점점 가까워지다가 1의 얼굴이 되었을 때 눈썹과 겹치게 되기 때문에 최댓값(1)에서 눈의 수직위치가 0.5 정도의 수직위치가 되도록 수정하였다. SAS macro faces 프로그램의 678번째 줄에서 $p\{6\}=p\{6\}*0.5를 추가하였다.$

Table 2.7에서 변수 7(눈썹의 곡률)은 2절에서 설명했듯이 얼굴의 인상을 결정하는 데에 영향을 미친다. 하지만 최솟값(0)의 경우 좋은 인상을 보이고 최댓값(1)의 경우 우울한 얼굴이다. 이 두 얼굴이 바뀌도

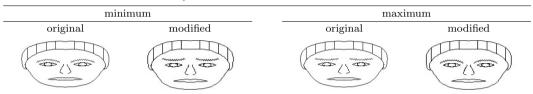
Table 2.6. Variable 6: vertical position of eyes



minimum: the vertical position of eyes is far from the nose.

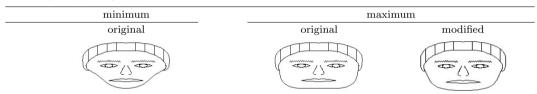
maximum: the vertical position of eyes is proper from the nose.

Table 2.7. Variable 7: curvature of eyebrows



minimum: the curvature of eyebrows looks downward. maximum: the curvature of eyebrows looks upward.

Table 2.8. Variable 13: face line



minimum: the face line is like an inverted triangle.

maximum: the face line is round.

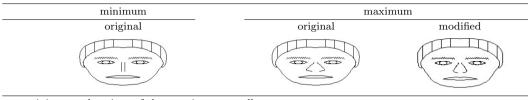
록 수정하였다. SAS macro faces 프로그램의 656번째 줄에서 $p\{7\} = 1 - p\{7\}$ 를 추가하였다.

Table 2.8에서 변수 13(얼굴 모양)의 최소값(0) 얼굴은 볼 살이 없어 야윈 얼굴을 보여준다. 기존의 최 댓값(1)에서의 얼굴은 사각형 모양으로 긍정적인 이미지 형성에 부족하기 때문에 가장 적절한 값을 가지는 0.7 정도의 얼굴 윤곽이 되도록 수정하여 둥글고 원만한 얼굴로 표현하였다. SAS macro faces 프로그램의 406번째 줄에서 $p\{13\} = p\{13\} * 0.7$ 를 추가하였다.

Table 2.9에서 변수 16(코의 모양)의 최댓값(1) 코 아랫 부분이 크게 벌어져서 얼굴의 전체적인 균형이 맞지 않다. 따라서 최댓값(1)의 코 모양이 0.7 정도의 코 모양이 되도록 수정하였다. SAS macro faces 프로그램의 597번째 줄에서 $p\{16\} = p\{16\}*0.7를 추가하였다.$

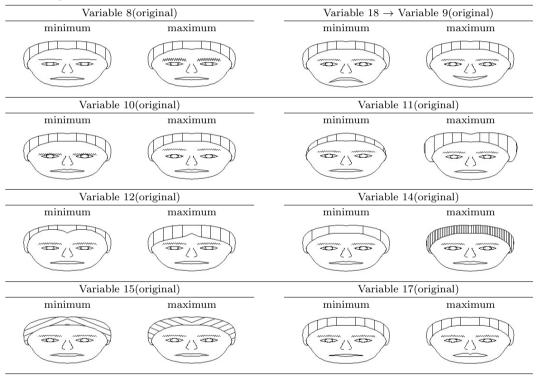
Table 2.10에는 수정하지 않은 변수들이다. 변수 8(눈썹의 밀도)는 최솟값(0)에서 최댓값(1)으로 값이 커질수록 눈썹은 점점 더 진해진다. 입의 곡률은 원 프로그램에서는 변수 18이었으나 수정 프로그램에서 변수 5와 변수 9가 변수 5로 합쳐졌기 때문에 변수 18이 변수 9로 변경되었다. 변수 9(입의 곡률)은 얼굴 이미지 형성에 가장 영향력 있는 변수이다. 최솟값(0)에서의 입은 화난 표정이며 최댓값(1)의 표정은 온화하다. 변수 10(눈썹의 수직위치)에서 최솟값(0) 눈썹의 수직위치는 눈과 거의 붙어 있으며 최댓값(1) 눈썹의 수직위치는 눈과 적절히 떨어져 있어 균형을 이룬다. 변수 11(윗머리의 모양), 변수

Table 2.9. Variable 16: nose line



minimum: the wings of the nose is very small. maximum: the wings of the nose is proper.

Table 2.10. Variable 8: density of eyebrows, Variable 9: curvature of mouth, Variable 10: vertical position of eyebrows, Variable 11: upper hair line, Variable 12: lower hair line, Variable 14: darkness of hair, Variable 15: hair shading slant, Variable 17: size of mouth



12(아랫머리의 모양), 변수 14(머리숱의 진한 정도)에서 최솟값(0)의 머리는 숱이 적으며 최댓값(1)의 머리는 숱이 많은 이미지로 표현된다. 변수 15(머리카락의 방향)에서 최솟값(0)의 머리카락은 아래 방향을 향하며 최댓값(1)의 머리카락은 윗 방향을 향한다. 변수 17(입의 크기)에서 최솟값(0)의 입의 크기는 작고 최댓값(1)의 입의 크기는 크다.

3. 얼굴 특징변수의 중요도

제 2절에서 Friendly (1992)의 SAS macro faces 프로그램에서 사용된 18개 얼굴 특징변수를 살펴보고

| variable | part of face | variable | part of face |
|----------|--|----------|-------------------------------|
| 1 | size of eyes | 10 | vertical position of eyebrows |
| 2 | size of pupils | 11 | upper hair line |
| 3 | position of pupils | 12 | lower hair line |
| 4 | slant of eyes | 13 | face line |
| 5 | horizontal position of eyes and eyebrows | 14 | darkness of hair |
| 6 | vertical position of eyes | 15 | hair shading slant |
| 7 | curvature of eyebrows | 16 | nose line |
| 8 | density of eyebrows | 17 | size of mouth |
| Q | curvature of mouth | | |

Table 3.1. Face feature variables

현 시대에 맞는 좋은 얼굴의 느낌에 따라 최종 17개 변수들의 수정안을 제시하였다. 즉, 원래 총 18개의 변수 중 변수 1~4, 변수 6~7, 변수 13, 변수 16을 수정하였고 변수 5와 변수 9는 통합하여 변수 5를 만 들었다. 또한 변수 18은 변수 9로 변경하였다. Table 3.1에는 수정한 얼굴 특징변수가 요약되어 있다.

체르노프 얼굴을 그리기 위한 수정된 SAS macro faces 프로그램에는 총 17개의 얼굴 변수가 존재하는데 다양한 시뮬레이션을 수행한 결과 얼굴의 인상을 결정하는데 중요한 역할을 하는 변수가 있는 반면에거의 영향을 주지 않는 변수도 있다. Table 3.2는 17개 변수마다 각 2개의 그림이 나열되어 있는데 각변수에서 첫 번째 그림은 해당하는 변수만 1이고 나머지 변수에 0을 준,두 번째 그림은 해당하는 변수만 0이고 나머지 변수에 1을 준 그림이다. Table 3.2를 통하여 어느 변수가 얼굴의 인상을 크게 좌우시키는지 알수 있다.

총 17개의 얼굴 특징변수 중에서 변수 9(입의 곡률)이 기쁜 인상 혹은 화난 인상을 만드는 데 큰 역할을 해 가장 중요성이 높았다. 또한 변수 13(얼굴의 모양)은 얼굴의 전체적인 느낌을 크게 변화시켰고, 변수 1(눈의 크기)도 얼굴 인상을 결정하는데 중요한 역할을 하였다. 변수 9(입의 곡률)과 더불어 일반적으로 얼굴 인상에 큰 영향을 끼친다고 생각되는 변수 7(눈썹의 곡률)이 별로 영향을 미치지 않음을 알 수 있다. SAS macro faces 프로그램에서는 눈썹의 모양이 곡선으로 처리되어 Figure 1.1에서 직선 모양의 눈썹보다는 더 온화해 보이기 때문이다.

변수 17(입의 크기)는 입의 크기에 따라서 웃거나 우는 표정을 더욱 강조되어 보이게 하는 역할을 하였지만 입의 곡률에 따라 얼굴 인상이 크게 달라지기 때문에 크게 중요하지 않는 것을 알 수 있다. 변수 11(윗머리의 모양), 변수 12(아랫머리 모양), 그리고 변수 14(머리카락의 진한 정도)는 서로 결합하여 일관성 있게 값이 커지거나 혹은 작아질 때는 영향력이 크나 서로 다른 방향으로 움직일 때의 영향력은 작아진다. 변수 2(눈동자 크기)는 눈 크기에 따라 얼굴 인상에 영향을 끼치고 어느 정도는 얼굴 인상을 결정하는데 중요하다고 판단된다. 변수 4(눈의 경사)는 눈의 크기가 작은 경우는 중요하였지만 눈의 크기가 큰 경우에는 최솟값과 최댓값 얼굴의 구별이 애매모호하기 때문에 중요성이 낮았다.

변수 5(눈과 눈썹의 수평위치), 변수 6(눈의 수직위치), 변수 10(눈썹의 수직위치)는 눈 및 눈썹의 위치만 변화시킬 뿐 얼굴의 인상에는 거의 영향을 끼치지 않았다. 그 밖에도 변수 3(눈동자의 위치), 변수 8(눈썹의 밀도)도 영향이 미비하여 중요성이 떨어졌다. 변수 15(머리카락의 방향)은 얼굴 인상을 결정함에 있어서 17개 변수 중에서 가장 영향을 끼치지 않는다는 것을 알 수 있다. 변수 16(코의 모양)은 최솟값과 최댓값의 얼굴 차이는 크지만 전체적으로 보았을 때는 얼굴 인상을 결정하는 데는 그렇게 중요한역할을 하지 않다고 판단된다. 이상과 같은 각 얼굴 특징변수의 중요도는 다변량 자료의 변수를 대응시키는 데 있어서 중요한 정보로 활용될 것이다. 그러나 이 또한 좋은 얼굴 이미지와 나쁜 얼굴 이미지에 대한 개인적 판단이 주관적이다는 지적을 피할 수는 없다.

Table 3.2. Importance of face feature variables



4. 자료분석

제 4절에서는 은행을 이용하는 고객들이 거래하고자 하는 은행을 선택할 때 고객의 입장에서 좋은 은행의 평가를 위하여 체르노프 얼굴분석을 수행하였다. 분석대상으로는 특수은행인 산업은행과 수협을 제외한 15개 은행으로 스탠다드차타드은행(구 제일은행), 경남은행, 광주은행, 국민은행, 기업은행(IBK), 농협, 대구은행, 부산은행, 신한은행, 씨티은행, 외환은행(KEB), 우리은행, 전북은행, 제주은행, 하나은행이다. 금융소비자연맹(http://www.kfco.org/)이 2013년 최초로 시행한 좋은 은행 평가를 위해 사용한 자료를 기초로 하여 은행의 안정성, 소비자 만족도, 건전성, 수익성을 나타낼 수 있는 11개 경제지표를 사용하였다. 또한 2012년 기준 금융감독원 전자공시시스템(http://dart.fss.or.kr/)의 공시자료와경영공시자료를 활용하였다. 소비자 만족도 부문에서 인지 및 신뢰도 지표는 금융소비자연맹이 2012년 11월 7,232명을 대상으로 한 설문조사 결과를 사용하였다. 본 연구에서 사용된 자료는 15행 11열의 다변량자료이다.

안정성을 나타내기 위한 경제지표로는 BIS자기자본비율(Bank for International Settlements capital adequacy ratio; BISCAR)과 유동성비율(liquidity ratio; LR)을 사용하였다. BIS자기자본비율은 국제 결정은행이 정한 은행의 위험자산(부실채권) 대비 자기자본비율로 각국 은행의 안정성 확보를 위해 최소 자기자본비율에 대해 국제적 기준을 마련한 것이다. BIS자기자본비율은 8% 이상으로 유지되기를 권고하고 있으며 본 연구의 분석 대상인 15개 은행은 모두 8%이상이며 씨티은행이 17.71%로 가장 높았고 기업은행이 12.37%로 가장 낮았다. 유동성 비율(원화 기준)은 은행이 자금수요에 있어서 얼마나 유동적으로 대처할 수 있는지를 판단하는 지표로써 이 값이 낮을수록 경영건전성 악화와 직결되기 때문에 금융감독원에서는 100% 이상을 유지하도록 규정하고 있다.

소비자 만족도를 나타내는 지표로는 고객 10만 명당 민원건수(number of consumer complaints per 100,000; NCC), 고객 10만 명당 민원증감률(rate of change in consumer complaints per 100,000; RCCC), 인지 및 신뢰도(perception and reliability; PR), 총자산 규모(total capital scale; TCS)를 사용하였다. 금융감독원에서는 금융회사의 자발적인 민원발생 예방 및 신속한 민원해결 노력을 유도하기위해 금융회사별 민원 발생건수를 정기적으로 공시하고 있다. 민원 발생건수가 적을수록 소비자 만족도가 높다. 고객 10만 명당 민원증감률은 금융감독원에서 조사한 전년대비 증감률을 의미하며 음의 값을 가질수록 서비스 향상이 이루어졌다고 판단한다.

건전성을 나타내는 지표로는 고정이하여신비율, 대손충당금적립률을 사용하였다. 고정이하여신비율(substandard credit ratio; SCR)은 총 여신에서 고정이하여신이 차지하는 비율로 재무 건전성을 의미한다. 일반적으로 은행 등 금융회사의 대출건전성을 나누는 기준은 5단계로 '정상', '요주의', '고 정', '회수의문', '추정손실'로 분류되는데 3개월 이상 연체된 여신을 '고정'이하여신 즉, 부실여신으로 추정한다. 고정이하여신비율이 높을수록 기업의 재무건전성이 위험하다는 의미로 고정이하여신비율이 높으면 정부의 대손충담금 적립 압박이 들어오므로 외부자금을 차입해야 할 필요성이 높아진다. 대손충당금 적립률(accumulation ratio of allowance for bad debts; ARABD)은 기업의 대출건전성을 판단할수 있는 지표로써 재무제표 상에 표시된 채권의 회수 불능으로 인해 생기는 기업의 손해와 비례한다. 대 손충당금 적립률이 커질수록 기업의 대출건전성을 의심해야 한다.

수익성을 나타내는 지표로는 총자산이익률, 순이자마진률, 당기순이익이 있다. 총자산이익률(return on asset; ROA)은 수익성을 판단할 수 있는 지표로서 기업의 총 자산대비 당기순이익의 값으로 일정기간 동안 얼마나 효율적으로 운용했는지를 알 수 있는 지표이며 금융회사의 가장 대표적인 수익성 지표로 총 자산이익률이 높을수록 기업의 수익성이 높으며 기업을 잘 운용하고 있다고 판단할 수 있다. 순이자마 진률(net interest margin; NIM)은 은행의 수익성을 판단할 수 있는 지표로써 순이자마진률은 순이자

| part of face | face feature variables | stability | consumer satisfaction | soundness | profitability |
|--------------|---------------------------------|-----------|-----------------------|-----------|---------------|
| head shape | var $11\sim12$, var $14\sim15$ | LR | TCS | SCR | NIM |
| eye | var $1\sim 8$, var 10 | LR | PR | SCR | NIM |
| nose | var 16 | 0.5 | 0.5 | 0.5 | 0.5 |
| mouth | var 9, var 17 | BISCAR | RCCC | ARABD | ROA |
| face shape | var 13 | BISCAR | NCC | ARABD | NIDT |

Table 4.1. Assignments of economic indicators by part to face feature variables

마진으로 은행에 발생하는 이익이기 때문에 순이자마진률 수치가 낮을수록 은행의 수익성이 낮은 것으로 판단할 수 있다. 당기순이익(net income during the term; NIDT)은 일정 기간 동안 기업이 창출한 수익으로 매출액에서 매출원가와 판매비, 관리비 등의 비용과 법인세를 제외한 순이익을 의미한다. 기업은 순이익으로 주주에게 배당을 하거나 사내유보로 자기자본을 축적한다.

SAS macro faces 프로그램에 다변량 자료를 입력할 때는 반드시 0과 1사이의 값을 입력해야하고 1의 값에 가까워질수록 좋은 얼굴을 의미하도록 변수를 조정해줄 필요가 있다. 따라서 고객 10만 명당 민원 건수, 고객 10만 명당 민원증감률, 그리고 고정이하여신비율은 작을수록 좋기 때문에 이 값들에 -1를 곱하여 순서를 미리 바꾸어 주어야 한다. 다음의 식은 각 변수의 값이 0부터 1까지의 실수값을 갖도록 조정해주는 식이다.

$$y_{ij} = \frac{x_{ij} - \min_j \{x_{ij}\}}{\max_j \{x_{ij}\} - \min_j \{x_{ij}\}}, \quad i = 1, 2, \dots, 15, \ j = 1, 2, \dots, 11,$$

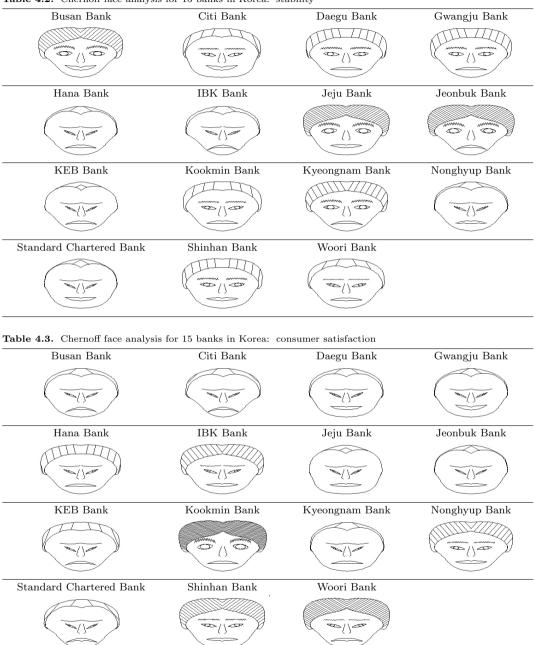
여기서 x_{ij} 는 i번째 은행에 대한 j번째 경제지표의 자료 값이다.

이제 안정성, 소비자 만족도, 건전성, 수익성의 각 부문별로 나누어 은행 간 비교를 한 후, 다시 4개 부문을 통합하여 종합적으로 은행 간 비교를 해보기로 하자. 안정성 부문은 2개, 소비자 만족도 부문은 4개, 건전성 부문은 2개, 수익성 부문은 각각 3개의 경제지표로 구성되어 각 부문에 속하는 경제지표 수는 얼굴 변수의 수 17개 보다 훨씬 적다. 이러한 경우 경제지표 수만큼 얼굴 변수를 대응시키고 나머지 얼굴변수들은 모두 0.5로 주는 방법이 있다. 그러나 그렇게 대응시킨 경우 2~4개 얼굴 부위만 미세하게 변화하여 얼굴의 특징이 잘 구별되지 않았다. 따라서 Table 3.1의 17개 얼굴 특징변수를 Table 4.1과 같이 최대 5개 그룹으로 나누어 각 부문별 경제지표에 대응시켰다. 즉, 변수 1~8, 변수 10을 눈 관련 변수 그룹, 변수 9 및 변수 17을 입 관련 변수 그룹, 변수 13을 얼굴모양, 변수 11~12, 변수 14~15를 머리모양 관련 변수 그룹으로 나누어 각 부문별 경제지표를 대응시켰고 코 모양을 나타내는 변수 16은 중 간값인 0.5의 값으로 고정시켰다. 코를 중심으로 머리, 눈, 입, 얼굴모양의 순으로 가까운 얼굴 부위끼리 묶어 경제지표를 대응시키므로서 얼굴 인상이 부위별로 흩어지지 않고 같은 이미지로 표현되도록 하였다.

Table 4.2는 안정성 부문에 관한 체르노프 얼굴분석을 보여준다. 안정성 지표 중 유동성비율(LR)을 머리 및 눈 관련 변수 그룹에, BIS자기자본비율(BISCAR)을 입 관련 및 얼굴모양 변수에 대응시켰다. 체르노프 얼굴분석 결과 제주은행과 부산은행이 가장 편안한 모습을 보이고, 씨티은행은 코 아랫쪽으로 좋은 모습이다. 실제로 제주은행은 BIS자기자본비율과 유동성비율이 15개 은행 중에서 각각 3번 째로 높았고, 부산은행은 유동성비율은 1위, BIS자기자본비율은 6위 였다. 반면에 씨티은행은 BIS자기자본비율은 1위이나 유동성비율이 9위이다. 매우 불만인 듯 보이는 은행은 기업은행(IBK), 하나은행, 외환은행(KEB), 농협 등이다.

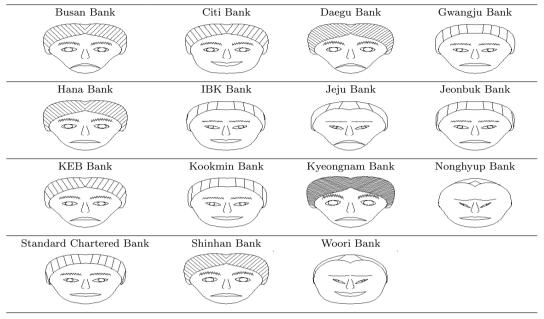
Table 4.3은 소비자 만족도 부문에 관한 체르노프 얼굴분석을 보여준다. 소비자 만족도 지표 중 총자산 규모(TCS)는 머리모양, 인지 및 신뢰도(PR)는 눈, 소비자 10만 명당 민원증감율(RCCC)은 입 관련 변

Table 4.2. Chernoff face analysis for 15 banks in Korea: stability



수 그룹에, 그리고 소비자 10만 명당 민원건수(NCC)는 얼굴모양에 각각 대응시켰다. 15개 은행 중에서 국민은행은 코 윗쪽 부위들이 가장 좋아보이는데 국민은행이 총자산 규모, 인지 및 신뢰도가 1위이

Table 4.4. Chernoff face analysis for 15 banks in Korea: soundness



다. 부산, 대구, 광주, 제주, 전북, 경남은행은 지역은행의 특성상 총자산 규모나 인지 및 신뢰도가 낮아 코 위쪽 부위들이 좋지 않으며 모두 비슷한 모습들이다. 광주은행은 코 아랫쪽 부위로는 가장 만족스러운 모습을 보여주는데 그것은 소비자 10만 명당 민원증감율과 소비자 10만명당 민원건수가 각각 1위와 4위이다. 씨티은행, 스탠다드차타스은행, 외환은행(KEB), 부산은행은 매우 불만스러운 얼굴모습을 보여준다.

Table 4.4는 건전성 부문에 관한 체르노프 얼굴분석을 보여준다. 건전성 지표 중 고정이하여신비율(SCR)을 머리모양 및 눈 관련 변수 그룹에, 대손충당금적립률(ARABD)을 입 관련 변수 그룹 및얼굴모양 변수에 대응시켰다. 체르노프 얼굴분석 결과 씨티은행이 가장 좋은 얼굴 모습을, 제주은행은 가장 안 좋은 얼굴 모습을 보여주었다.

Table 4.5는 수익성 부문에 관한 체르노프 얼굴분석을 보여준다. 수익성 지표 중 순이자마진율(NIM)을 머리모양 및 눈 관련 변수 그룹에, 총자산이익률(ROA)을 입 관련 변수 그룹에, 당기순이익(NIDT)을 얼굴모양 변수에 각각 대응시켰다. 체르노프 얼굴분석 결과 부산 및 대구은행이 가장 좋은 얼굴을 보여준다. 부산, 대구, 광주, 제주, 전북, 경남은행은 지역은행의 특성상 당기순이익 절대 액수가 작기 때문에 얼굴 하관이 좁은 모습을 보여주었다. 하나은행과 농협은 전반적으로 가장 좋지 않은 모습을 보여주고 있다.

마지막으로 4개 부문을 통합하여 총 11개의 경제지표를 모두 사용하여 은행 간 비교를 하였다. Table 4.6과 같이 얼굴 특징변수와 은행경제지표를 대응시켰다. 변수 9(입의 곡률)이 얼굴 인상을 크게 결정하기 때문에 고객 입장에서 좋은 은행 평가에 가장 중요한 안정성 분야의 BIS자기자본비율(BISCAR)을 대응시켰다. 또 다른 안정성 분야의 지표인 유동성비율(LR)을 변수 7(눈썹의 곡률)에 대입하여 안정성이 얼굴의 전체적인 이미지를 구성할 수 있도록 하였다. 소비자 만족도로 얼굴 모양을 결정하기 위해서 총자산 규모(TCS)를 변수 14(머리숱의 진한 정도), 소비자 10만 명당 민원건수(NCC)를 변수 12(아랫

Table 4.5. Chernoff face analysis for 15 banks in Korea: profitability

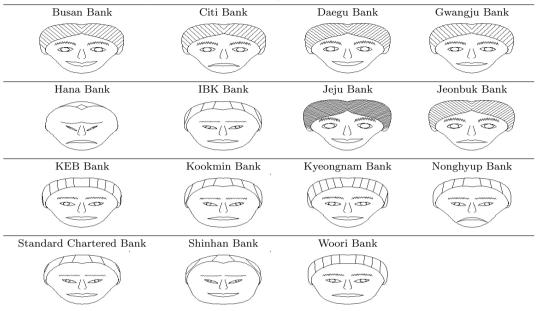
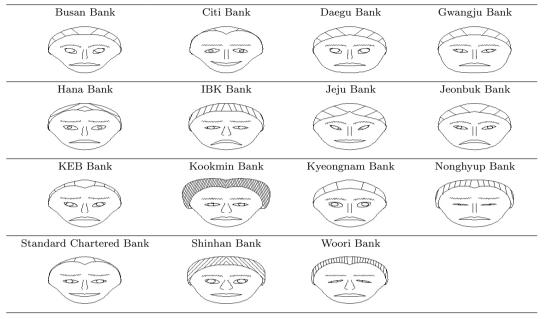


Table 4.6. Assignments of economic indicators to face feature variables

| face feature variable | part of face | economic indicator | |
|-----------------------|--|--------------------|--|
| 1 | size of eyes | SCR | |
| 2 size of pupils | | ROA | |
| 3 | position of pupils | 0.5 | |
| 4 | slant of eyes | ARABD | |
| 5 | horizontal position of eyes and eyebrows | 0.5 | |
| 6 | vertical position of eyes | 0.5 | |
| 7 | curvature of eyebrows | $_{ m LR}$ | |
| 8 | density of eyebrows | 0.5 | |
| 9 | curvature of mouth | BISCAR | |
| 10 | 10 vertical position of eyebrows | | |
| 11 | upper hair line | PR | |
| 12 | lower hair line | NCC | |
| 13 | face line | RCCC | |
| 14 | darkness of hair | TCS | |
| 15 | hair shading slant | NIM | |
| 16 | nose line | NIDT | |
| 17 | size of mouth | 0.5 | |

머리 모양), 인지 및 신뢰도(PR)를 변수 11(윗머리 모양)에 대응시켰고, 특히 고객 10만 명당 민원증감률(RCCC)을 변수 13(얼굴 모양)에 대응시켜 고객의 만족도를 나타내게 하였다. 건전성을 나타내는 지표인 고정이하여신비율(SCR)을 변수 1(눈의 크기)에 대응시키고 대손충당금적립률(ARABD)을 변수 4(눈의 경사)에 대응시켜 눈의 이미지를 표현하였다. 또한, 총자산이익률(ROA), 순이자마진율(NIM),

Table 4.7. Chernoff face analysis for 15 banks in Korea



당기순이익(NIDT)과 같은 수익성지표는 고객 입장의 좋은 은행 평가에 있어서는 상대적으로 작은 비중을 차지하여 각각 변수 2(눈동자 크기), 변수 15(머리카락의 방향), 변수 16(코 모양)에 대응시켰다. 나머지 6개의 얼굴 특징변수는 동일하게 <math>0.5의 값을 주었다.

Table 4.7은 국내 15개 은행에 대해 소비자 관점에서 종합적으로 평가한 체르노프 얼굴분석 결과를 보여준다. 신한은행이 전체적으로 가장 균형 있고 좋은 얼굴을 보여주며, 기업은행(IBK), 외환은행(KEB), 하나은행, 전북은행, 경남은행이 화난 인상을 보여주었다. 금융소비자연맹에서 공시한 은행별 종합 순위는 안정성, 소비자 만족도, 건전성, 수익성에 각각 40%, 30%, 20%, 10%의 가중치를 주어산출하였다. 금융소비자연맹의 은행평가 순위 결과표에서는 신한은행이 안정성, 소비자 만족도, 건전성, 수익성의 각 분야에서 모두 3위 내에들어 종합순위에서 1위를 차지하였다. 종합순위 2위를 차지한씨티은행의 경우얼굴 표정은 웃는 인상이나얼굴 모양이 좋아 보이지 않다. 얼굴 모양, 머리 모양등을나타내는얼굴의 전체적인들은 소비자 만족도를 나타내는데 소비자 만족도 부문에서 최하위를 기록하였다. 반면에 국민은행의 경우 머리의 모양이 눈에 띄게 풍성해 보인다. 국민은행은 고객인지 및 신뢰도와 총자산 규모에서 1위를 하였고, 10만명당 민원건수와고객 10만명당 민원증감률에서 각각 10위, 11위를하였다.

5. 결론

Chernoff (1973) 이후의 체르노프 얼굴분석에 관한 연구를 추적하면서 이 방법론이 대중적으로 잘 쓰이지 않았고 최근의 논문도 거의 찾기 힘들었다. 그것은 얼굴의 인상을 좌우하는 얼굴 특징변수에 대응시키는 다변량 변수의 선택이 매우 주관적이며 그 결과로서 표현되는 얼굴 인상에 대한 정서적 해석이 민감할 수 있기 때문으로 판단된다. 그러나 체르노프 얼굴분석은 일반인도 빠르고 쉽게 얼굴 이미지를 통하여 각 관측치의 군집화 혹은 다변량 관측치에 대한 종합적인 판단을 가능하게 하므로 매우 흥미로운

분석법이다. 본 연구에서는 SAS macro faces 프로그램에서 사용되는 18개 얼굴 특징변수를 현 시대에 맞는 좋은 얼굴의 관점에서 수정안을 제시하였다. 또한 수정된 프로그램을 이용한 체르노프 얼굴분석을 통하여 국내 15개 은행에 대해 고객의 관점에서 보는 좋은 은행평가를 시도하였다. 실제 자료분석을 통하여 체르노프 얼굴분석에 대한 근본적인 한계를 느낄 수 있었다. 얼굴의 인상 선호도에 있어서 객관적인 판단이 어려워 각 은행들의 체르노프 얼굴에 대해 주관적으로 느낄 수 있기 때문에 자료에 대한 추가적인 통계분석이 참고 되어야만 체르노프 얼굴분석이 가지고 있는 효과를 발휘할 수 있었다. 또한 얼굴특징변수에 어떤 다변량 변수를 대응시키느냐에 따라 다른 결과가 나올 수 있는 점은 가장 큰 어려움이라고 할 수 있다.

References

- Anderson, E. (1960). A semi-graphical method for the analysis of complex problems, *Technometrics*, 2, 287–292.
- Andrews, D. F. (1972). Plots of high-dimensional data, Biometrics, 28, 125-136
- Apaiwongse, T. S. (1995). Facial display of environmental policy uncertainty, Journal of Business and Psychology, 10, 65–74.
- Astel, A., Astel, K., Biziuk, M. and Namieśnik, J. (2006). Classification of drinking water samples using the Chernoff's faces visualization approach, *Polish Journal of Environmental Studies*, **15**, 691–697.
- Bertin, T. (1967). Semiologie Graphique, Paris: Gauthier-Villars.
- Chernoff, H. (1973). The use of faces to represent points in k-dimensional space graphically, *Journal of the American Statistical Association*, **68**, 361–368.
- Friendly, M. (1992). SAS Macro Program: faces, Available at http://www.datavis.ca/sasmac/faces.html
- Golden, L. L. and Sirdesai, M. (1992). Chernoff: A useful technique for comparative image analysis and representation. Advances in Consumer Research, 19, 123–128.
- Goldwyn, R. M., Friedman, H. P. and Siegel, T. H. (1971). Iteration and interaction in computed data bank analysis; Case study in physiological classification and assessment of the critically ill, Computers in Biomedical Research, 4, 607–622.
- Huff, D. L., Mahajan, V. and Black, W. C. (1981). Facial representation of multivariate data, The Journal of Marketing, 45, 53-59.
- Kruskal, J. B. and Wish, M. (1978). *Multidimensional Scaling*, Sage University Paper Series on quantitative Applications in the Social Sciences, 07-011, Sage Publications, Beverly Hills and London.
- Lott, J. A. and Durbridge, T. C. (1990). Use of Chernoff faces to follow trends in laboratory data. *Journal of Clinical Laboratory Analysis*, 4, 59-63.
- Spinelli, J. G. and Zhou Y. (2004). Mapping quality of life with chernoff, ESRI conference paper, Available at http://gis.esri.com/library/userconf/educ04/papers/pap5000.pdf
- Turner, E. (1979). Life in Los Angeles 1970, Available at http://www.pinterest.com/pin/98727416802725604/

SAS macro faces를 사용한 체르노프 얼굴 분석에 의한 좋은 은행 평가

0]정은 a · 정혜선 a · 김민지 a · 김지현 a · 소영숙 a,1

^a전남대학교 통계학과

(2013년 09월 27일 접수, 2013년 11월 09일 수정, 2013년 11월 25일 채택)

요 약

Friendly (1992)의 SAS macro faces 프로그램은 다변량 시각화 표현법 중 하나인 체르노프 얼굴 분석을 위한 프로그램이다. 본 연구에서는 faces에서 사용된 18개 얼굴 특징 변수를 살펴보고 현 시대의 흐름에 맞는 좋은 얼굴의 정의에 따라 수정된 프로그램을 제시하였다. 또한 은행의 안정성, 소비자 만족도, 건전성, 수익성을 나타내는 11개 은행경제지표에 기초한 체르노프 얼굴분석을 통하여 국내 15개 은행에 대해 소비자 입장에서 좋은 은행 평가를 수행하였다.

주요용어: 체르노프 얼굴, SAS macro faces 프로그램, 좋은 은행 평가.

이 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임 (NRF-2011-0022864)

¹교신저자: (550-757) 광주광역시 북구 용봉동 300번지, 전남대학교 통계학과, 교수. E-mail: ysson@jnu.ac.kr