

한글 글꼴 유사성 판단을 위한 획 요소 속성의 영향력 분석

박동연[†], 전자연^{**}, 임서영^{***}, 임순범^{****}

A Study on Influence of Stroke Element Properties to find Hangul Typeface Similarity

Dong-Yeon Park[†], Ja-Yeon Jeon^{**}, Seo-Young Lim^{***}, Soon-Bum Lim^{****}

ABSTRACT

As various styles of fonts were used, there were problems such as output errors due to uninstalled fonts and difficulty in font recognition. To solve these problems, research on font recognition and recommendation were actively conducted. However, Hangul font research remains at the basic level. Therefore, in order to automate the comparison on Hangul font similarity in the future, we analyze the influence of each stroke element property. First, we select seven representative properties based on Hangul stroke shape elements. Second, we design a calculation model to compare similarity between fonts. Third, we analyze the effect of each stroke element through the cosine similarity between the user's evaluation and the results of the model. As a result, there was no significant difference in the individual effect of each representative property. Also, the more accurate similarity comparison was possible when many representative properties were used.

Key words: Characteristics Properties of Hangul Shape, Hangul Font Similarity, Font Similarity Calculation Model, Stroke Element Property

1. 서 론

전자 출판이 성장하는 것은 물론, 웹 페이지 제작, 파워포인트(PowerPoint) 등 다양한 발표 자료 제작과 같은 콘텐츠 산업이 부상함에 따라 콘텐츠의 시각적인 영향력을 보다 효과적으로 전달하기 위해 다양한 스타일의 글꼴이 사용되고 있다. 그에 따라 폰트 산업도 함께 활발해지고 있으며, 현재 사용되고 있는

한글 글꼴은 육 천종이 넘게 존재한다[1]. 이렇게 다양한 글꼴들이 사용됨에 따라 여러 문제도 함께 발생하고 있다. 두드러지는 문제로는 콘텐츠 내 폰트가 파일 및 콘텐츠를 전달받은 사용자의 컴퓨터 내에 설치되지 않았을 경우가 있다. 예를 들어, 사용자 간 워드(Word), 파워포인트 등의 문서 파일을 공유할 때, 파일 내에는 사용되었지만 사용자의 컴퓨터에는 설치되지 않은 폰트로 인해 출력 레이아웃이 엉클어

※ Corresponding Author : Soon-Bum Lim, Address: (04 310) Cheongpa-ro 47-gil 100, Yongsan-gu, Seoul, Koera, TEL : +82-2-710-9424, FAX : +82-2-710-9704, E-mail : sblim@sm.ac.kr

Receipt date : Sep. 4, 2020, Revision date : Nov. 9, 2020
Approval date : Nov. 13, 2020

[†] Dept. of IT Engineering, SookMyung Women's University (E-mail : yeon0729@sm.ac.kr)

^{**} Dept. of IT Engineering, Graduate School, SookMyung Women's University (E-mail : jyjeon@sm.ac.kr)

^{***} Dept. of IT Engineering, SookMyung Women's University (E-mail : lsy783@sm.ac.kr)

^{****} Dept. of IT Engineering and Research Institute ICT Convergence, Sookmyung Women's University

※ 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-2018R1A4A1025559) and Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MS IP) (No.2016-0-00166)

지는 등의 문제가 발생한다. 또 다른 문제로는 구체적인 폰트 정보 파악이 어렵다는 것이다. 이미 수많은 폰트가 존재하기 때문에 비전문가의 입장에서는 다른 사람의 콘텐츠 내 글자 이미지만으로는 구체적인 폰트 정보를 알기 어렵다. 이 외에도 유사 폰트 검색이 어렵다는 문제가 있다. 앞선 경우에서 원하는 폰트를 찾더라도 저작권 문제 등으로 인해 사용이 불가능할 수도 있다. 따라서 유사한 스타일의 다른 폰트를 찾고자 하여도, 비전문가의 입장에서는 폰트 스타일을 파악하고 검색하는 것도 어려운 일이다.

이러한 문제들을 해결하기 위해 폰트 인식 및 유사 폰트 추천에 관하여 영문 폰트와 중문 폰트를 중심으로 다양한 연구가 진행되었다. 대표적으로 문서 전반 이미지의 폰트 텍스처 정보를 이용하거나[2], 단일 글자 이미지를 통한[3] 폰트 인식 알고리즘 연구가 있다. 이 외에도 최근에는 딥러닝 기술을 접목한 다양한 연구들이 진행되었는데, CNN 모델을 이용하여 유사 글꼴을 식별하거나[4] kNN 알고리즘을 이용한 유사 글꼴 그룹핑[5], RNN 모델에 기반한 글꼴 분석 및 검색에 대한 연구[6] 등이 있다. 반면 한글 폰트는 구성이 복잡하고 어려울 뿐더러 한글 글꼴 검색 서비스는 매우 초보적인 단계로써 한글 폰트를 검색하고 비교 분석하는 연구는 수동적인 방법으로 진행되었다[7,8].s

향후 연구에서 한글 글꼴 내 유사 폰트 추천에 대한 자동화를 이룩하기 위한 기초 자료로써, 본 연구에서는 획 모양 요소별로 폰트 유사성 판단에 미치는 영향력을 분석하고자 하였다. 기존 영문에 대한 연구에서는 글자 이미지 단위로 비교를 진행하였지만, 한글의 경우에는 구조가 복잡하여 글자 단위의 비교로는 불충분하다. 따라서 한글 구조와 특징을 고려한 획 모양 요소를 기반으로 하여 폰트 간 유사성을 판단할 수 있는 모델을 구축하고자 하였다.

이를 위해 선행 연구[9]에서 제안된 여러 글꼴 속성 중 한글 구조의 획 모양 요소에 기반을 둔 대표적인 핵심 속성들을 선정한 뒤 측정 방법을 정의하였다. 선정한 속성 값을 이용하여 폰트 간 특징 거리를 계산하고, 이를 기반으로 폰트 간 유사성 판단이 가능하도록 하는 계산 모델을 설계하였다. 그리고 대표 폰트에 대한 실제 사용자들의 유사성 평가 결과와 본 연구에서 설계한 모델의 유사성 평가 결과의 일치 정도를 계산하여, 획 요소별 글꼴 유사성 판단에 미

치는 영향력을 분석하였다.

본 논문의 구성은 다음과 같다. 2장에서 선행 연구의 한글 글꼴 분석에 사용된 PANOSE[10]와 선행 연구[9]에서 제안한 글꼴 속성에 관해 설명한다. 그리고 한글 구조를 기반으로 연구를 진행하기 위해 한글 글꼴의 계층적 구조 정보에 관해서 설명하고, 연구의 발전 방향성 모색을 위한 딥러닝 기술을 접목한 폰트 연구를 알아보도록 한다. 3장에서는 선행 연구에서 제시한 한글 글꼴 속성 중 획 모양 요소 기반의 핵심 속성을 선정한다. 그리고 속성값을 이용하여 유사성을 파악할 수 있는 개별 획 요소 기반 폰트 유사성 계산 모델을 설계한 후, 실험을 위한 대표 폰트 13종을 선정한다. 4장에서는 대표 폰트에 대한 유사성 순위를 사용자들로부터 설문을 진행하여 도출하고, 본 논문에서 제시한 모델의 유사성 순위 결과와의 일치 정도를 코사인 유사도를 통해 계산한다. 그리고 분산 분석을 이용하여 속성별 영향력을 파악한다. 마지막 6장에서는 최종적인 결론과 향후 연구 방향에 대해서 제시한다.

2. 관련연구

2.1 PANOSE[10]

PANOSE 시스템은 10가지의 시각적 특성만으로 폰트를 분류하여 이를 기반으로 시스템에 존재하지 않는 폰트를 대체하기 위한 폰트를 찾을 때 주로 사용된다. 크로스 플랫폼 환경에서 누락된 폰트 및 호환되지 않는 폰트로 인해 발생하는 문제를 해결해 줌으로써, 이러한 문서 표현의 유연성 덕분에 World Wide Web에서 그 장점을 크게 발휘한다. 현재는 HP가 해당 기술을 소유하고 있고, W3이 관리 및 표준화를 하였다.

PANOSE 시스템의 핵심은 Typeface Classification Number로 대비, 무게, 세리프 스타일과 같은 글꼴의 중요한 시각적 특성을 10바이트로 간결하게 설명할 수 있다는 것이다. 10가지 속성은 다음 그림 Fig. 1과 같다. 다양한 글꼴 공급 업체에서 글꼴 PANOSE 번호를 측정 및 할당하는 데에는 객관적인 기준 및 기술이 필요한데, Hewlett Packard의 PANOSE 분류 가이드인 "The Gray Book"가 그 기준이 된다. 현재 PANOSE를 사용하여 유사 글꼴을 찾는 방법은 PANOSE Mapper 소프트웨어를 사용하는 것이다.

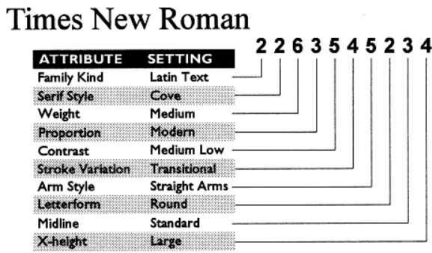


Fig. 1. PANOSE number of Times New Roman fonts[10].

PANOSE Mapper 소프트웨어는 요청된 폰트와 사용 가능한 폰트의 PANOSE 번호를 비교하여 주어진 시스템에서 가장 폰트인지를 파악한다. 그리고 개별 PANOSE 번호를 비교하고, 인쇄상의 중요성에 따라 가중치를 부여하고 합산하여 숫자로 된 시각적 거리를 제공한다.

2.2 한글 모양 속성 기반 글꼴 추천 시스템[9]

해외에서는 다양한 기술의 발전을 기반으로 폰트 전문가가 아닌 일반 사용자들도 쉽게 접근 가능한 이미지 기반 유사 폰트 찾기[11] 및 폰트 페어링[12] 서비스를 제공 중이다. 하지만 서비스 대부분은 영문 폰트를 기준으로 구축되어으며 한글 폰트에 대해서는 서비스를 제공하지 않는다. 심지어 국내에서 이와 같은 서비스를 제공하는 곳은 찾기 어려운 실정이다.

따라서 선행 연구[9]에서는 PANOSE를 기반으로 한글 모양 분류 체계를 구축하여 ‘굵기 관련’, ‘기울기’, ‘홀자, 닿자’, ‘간격, 공간’, ‘돌기/부리’, ‘굴림, 접촉’, ‘뿔음’, ‘윤곽’, ‘재질’의 9개의 그룹으로 분류 가능한 75개의 한글 글꼴 특성을 제시하였다. Fig. 2는 선행연구에서 제작한 감성 분류 체계와 TTA 모양 분류 체계 기반의 한글 글꼴 추천 시스템의 인터페이스 화면이다.

2.3 한글 글꼴의 계층적 구조 정보[13]

한글 폰트 추천과 같은 한글 글꼴 관련 연구와 서비스가 미비하거나 없는 것에는 한글 자체가 복잡한 글자 구조로 구성이 되어있어 분석이 어려울뿐더러, 공통으로 사용할 수 있도록 정립된 기준이 없기 때문이다. 따라서 한글 글꼴 정보를 Fig. 3와 같은 계층적인 형태로 모델링하고, 획 요소를 기반으로 글꼴 별 특징 속성 정보를 어떻게 담아낼 수 있는지 구조를 파악하고 지도 역할을 할 구조맵을 그리는 연구[13]



Fig. 2. Font recommendation system based on the sense classification system and the TTA shape classification system[9].

가 진행되었다. Fig. 3에서 확인할 수 있는 한글 글꼴의 계층적 구조 정보는 3가지 층으로 구성되어 있다. 우선 한글의 제작 원리에 따른 자모음의 상형성을 기본으로 하여 한글의 구조 정보를 설계한 기본구조 수준이 가장 낮은 층이다. 이에 폰트의 특성을 담아내는 속성이 증가할수록 상위 계층으로 올라가게 된다. 기본 구조 수준에서 설계된 구조를 바탕으로 모든 폰트에 공통으로 적용될 수 있도록 설계된 것이 두 번째 층인 정형 폰트 수준이다. 마지막으로 돌출체 및 바탕체처럼 정형화된 모양이 아닌 감각적이고 예술적인 모양을 포함하여 설계된 것이 마지막 최상위 층인 비정형 폰트 수준이다.

이때 기본구조 수준에서 설계한 한글의 구조 정보를 담고 있는 구조맵(Structure Map)은 Fig. 4와 같이 Node와 Link로 구성되어 있다. Fig. 4에서는 구조맵을 구성하는 Node와 Link의 세부적인 종류와 그에 대한 표기법 등의 설명을 확인할 수 있다. 이를 기반으로 한글 폰트를 구성하는 획 요소는 17개로 정리할 수 있다.

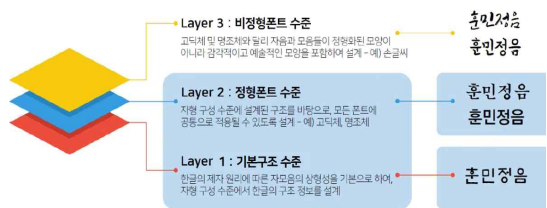


Fig. 3. Hierarchical structure information of Hangul fonts [13]

명칭	분류	표기법	설명
Horizontal Vertical	Node	H_u, H_d, V_u, V_d	기본적인 가로, 세로 획을 지칭. 시작(start)과 끝(end)으로 나뉨
Turn	Node	T_r, T_l	획에서의 방향을 지칭. 오른쪽(right)과 아래(down)으로 나뉨
Left Right	Node	L, R	'ㄱ' 'ㅋ' 'ㆁ'에서 좌우로 빠져나가는 획을 지칭
O	Node	O	'ㅇ'에서 끝맺음으로 나타나는 이용을 지칭
Cap	Node	C_u, C_d	'ㄷ'에서 끝맺음으로 나타나는 첫 부분의 받침 혹은 돌기 등을 지칭. 시작(start)과 끝(end)으로 나뉨
Virtual Node	Node	●	획과 다시 링크(Branch Link)를 잇기 위한 가상적 노드
Normal Link	Link	—	노드와 노드 사이를 잇는 일반 연결(링크)
Branch	Link	B (Branch)	링크와 링크 혹은 노드와 링크 사이의 연결을 지칭
Joint	Link	J (Joint)	노드와 노드가 겹쳐져서 나타날 때의 연결을 지칭

Fig. 4. Structure map of basic structure floor of Hangul fonts[13].

2.4 한글 글꼴과 딥러닝

다양한 기술들이 빠른 속도로 발전하고 기술에 대한 접근성이 더 쉬워짐에 따라 글꼴 분야에서도 여러 연구가 진행되고 있다. 최근에 다양한 분야에서 활발히 사용되고 있는 딥러닝 기술을 글꼴에도 접목한 연구들이 진행되었는데, 앞서 언급한 이미지를 통한 유사 글꼴 식별[4] 및 그룹핑[5], 글꼴 분석 및 검색[6]과 생성 학습을 통한 대규모 태그 기반 글꼴 검색[14] 등이 있다. 국내에서는 한글 폰트 이미지 분류를 위한 클래스 세트 찾기[15], 손글씨 이미지를 같은 스타일의 다른 글자 이미지로 변환[16], 글자 이미지 간 변환을 위한 대표 글자 쌍 평가[17] 등이 진행된 바가 있다.

이 외에도 R-CNN을 사용하여 한글 획 요소를 학습시켜, 자동으로 서로 다른 폰트에서도 같은 획 요소를 찾아주는 연구[18]가 진행되었다. 앞서 언급하였듯이 한글 글꼴 시스템은 단순하게 글꼴명 혹은 글꼴 제작사명만으로 글꼴 정보를 저장하여 수동적으로 글꼴 정보를 제공[7,8]하므로 수많은 한글 글꼴의 모든 모양 특성을 이해하고 사용하기는 어렵다. 따라서 한글 글꼴 전반적으로 폰트의 특징을 대표할 수 있는 글꼴 특징 요소를 선정하고, 이를 자동 추출할 수 있도록 하여 글꼴 분석을 더 쉽게 하고자 하였다. Fig. 5는 위 연구[18]에서 선정한 대표 글꼴 특징 요소이다.

획 요소 검출을 위해 170종의 글꼴마다 Fig. 5에서 볼 수 있는 8개의 획요소를 선별하여 총 1,360장의 데이터 세트로 학습을 진행하였다. 학습 모델로는 F aster R-CNN inception v2를 사용하였고, 획 요소별로는 평균 43,109번씩 학습을 진행하였다. 실행 데이터 세트를 위해 다양한 계열로 선정한 13종의 글꼴에서도 8개의 모든 획 요소를 검출하도록 104장의 글자

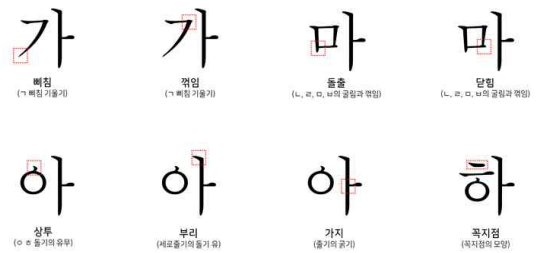


Fig. 5. Characteristics of stroke elements in representative Hangul fonts for similarity[18]

이미지 세트를 구성하였다. 이후 앞서 학습시킨 모델을 사용하여 획 요소 검출을 진행하였을 때, 104장의 글자 이미지에서 모두 정의한대로 104개의 획 요소가 검출되는 100%의 성공률을 확인할 수 있었다. 결과 이미지는 Fig. 6과 같다. 이전에는 글꼴 분석 및 비교를 위해서 일일이 사람이 획 요소를 직접 찾아야 했다면, 현재는 해당 연구를 통해 이러한 반복적인 부분이 자동화가 가능해짐으로써 사람의 개입으로 소모되는 다양한 시간적, 경제적 자원을 아낄 수 있게 되었다. 본 연구와 글꼴 특징 요소 자동 추출 연구를 서로 접목한다면 획 요소별 핵심 속성값 자동 측정과 다양한 폰트 간 유사성 비교까지 자동화가 가능해질 것이다.

	Serif			Sans-Serif		
	HY신명조	HY경명조	산돌독수리	돋움체	맑은고딕	타이포도당
bkchilm	가	가	가	가	가	가
kkeokilm	가	가	가	가	가	가
dokhul	마	마	마	마	마	마
dschilm	머	머	머	머	머	머
sangtu	하	하	하	하	하	하
huri	아	아	아	아	아	아
gaji	아	아	아	아	아	아
kkokijjum	하	하	하	하	하	하

Fig. 6. Result of stroke element detection training[18].

3. 핵심 속성 선정 및 유사성 측정 모델 개발

3.1 개요

본 연구는 다음과 같은 순서로 진행된다. 선행 연구[9]에서 정리한 75개의 한글 글꼴 속성 중 폰트 유사성 비교에 유효하게 영향을 미칠 것으로 판단되는 핵심 속성을 한글의 획 모양 요소를 기반으로 선정한다. 그리고 선정된 핵심 속성의 속성값을 측정하는 방법을 정의한다. 이를 기반으로 폰트 간 유사성을 측정 및 비교할 수 있는 유사성 측정 모델을 개발한 뒤, 모델의 실험을 위한 글꼴 표본인 대표 폰트를 선정한다. 선정된 대표 폰트의 핵심 속성값을 측정하고, 앞서 개발한 유사성 측정 모델을 사용하여 대표 폰트 간 유사성 순위를 계산한다. 위 과정을 그림으로 나타냈을 때 Fig. 7과 같다.

3.2 핵심 속성 선정

선행 연구[9]에서 도출한 한글 폰트 속성 75개는 한글 글자 구조를 이루는 획요소(Stroke Element)와 골격(Skeleton), 공간 정보(Spacing) 3가지 모두를 고려한 속성 목록이므로, 이 중 본 연구 목적에 적합한 것으로 판단되는 핵심 속성만을 다시 선정하고자 하였다[18]. 핵심 속성 선정을 위한 평가 작업은 총 2차에 걸쳐, 아래의 두 가지의 평가 척도를 기반으로 진행되었다. 평가자로서는 폰트에 관련하여 전문적

인 지식과 경험을 보유하고 있는 폰트 업계 종사자 5명이 참여하여 핵심 속성 선정을 위한 평가를 실시하였다.

- [1차 평가 척도] 이미지 분석 및 파라미터로 측정할 수 있는 특성인가?
- [2차 평가 척도] 유사성 분석을 할 때 있어서 중요하고 유의미한 특성인가?

1차 평가에서는 ‘이미지 분석 및 파라미터로 측정할 수 있는 특성인가?’를 기준으로 정하였고, 그에 따라 75개의 속성 목록에서 25개의 속성들이 제외되어 50개의 속성 목록을 도출할 수 있었다. 그에 이어 2차 평가에서는 앞서 도출한 50개의 속성 목록에 대해 ‘유사성 분석을 할 때에 있어서 중요하고 유의미한 특성인가?’를 기준으로 5점 리커트 척도를 사용하여 각 항목 당 중요도 점수를 매기도록 하였다. 그 결과는 Table 1과 같다. 4점 이상은 9개, 3점 이상 4점 미만은 20개, 2점 이상 3점 미만은 17개, 1점 이상 2점 미만은 4개로 확인할 수 있었다. 최종적으로 3점 이상의 항목 중에서 한글 구조의 획 모양 요소를 기반으로 구성된 7개의 핵심 속성 목록을 도출하였다.

3.3 핵심 속성 값 측정

핵심 속성 목록을 선정된 후에는 해당 속성의 값을 어떻게 측정할지 측정 기준을 정의해야 정확한 속성 값 측정 및 유사성 비교 모델 구축이 가능하다. 따라서 속성별로 속성값의 측정 기준을 정의하였는데, 이때 속성값을 절대적인 수치값이 아닌 비슷한 형태와 모양을 띤 속성들을 묶음 클래스 분류 값을 의미하는 것으로 하였다. 이때 바탕체와 같은 Serif 계열의 글꼴과 San-Serif 계열의 글꼴에 대한 분리를 명확하게 하기 위해, 해당 정보를 반영하는 단계에서는 다른 분류 단계보다 클래스 분류 값을 멀게 지정하였다. 자세한 클래스 분류는 다음 Fig. 8과 같다.

첫 번째 핵심 속성은 상투 획 요소에 위치한 상투의 꺾임 각도이다. 둥글 이용의 획 영역과 돌기 획 요소가 만나는 지점의 각도를 4단계(0-2-3-4)로 분류하였다. 두 번째는 부리 획 요소에 위치한 세로 줄기의 부리 각도이다. 세로 줄기에서 부리의 유무, 줄기와 부리의 연결 부분의 외각도를 기준으로 3단계(0-2-4)로 분류하였다. 세 번째는 가지 획 요소에 위치한 줄기의 굽기이다. ‘KopubWorld 돋움 Medium’

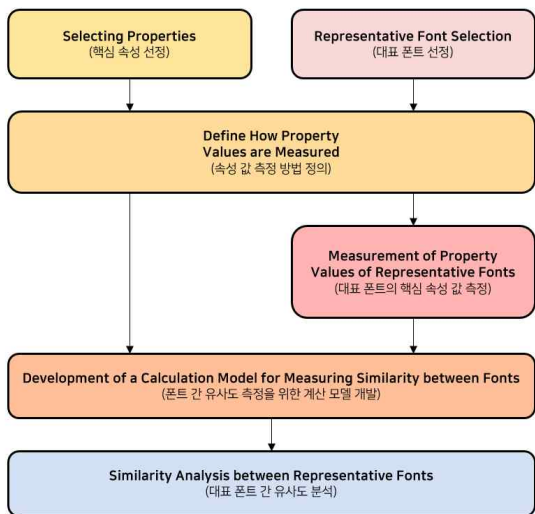


Fig. 7. Flow chart of property selection and similarity calculation model construction.

Table 1. The importance evaluation score of stroke elements and properties of Hangeul fonts

No	Property	Score	No	Property	Score
Thickness			38)	‘ㅇ,ㅎ’ symmetry	3.25
1)	two point thickness of transverse stem	3	39)	vertical letter-line straight	2.75
2)	thickness of stem	4.25	40)	character width	3
3)	stem thickness uniformity	3.75	42)	gap between visual midline and bottom line ☆	4.25
4)	uniformity of ‘ㅇ’ thickness	2.75	45)	proportional width or fixed width	3.5
5)	‘ㅍㅍichim’ thickness uniformity	3	46)	jongseong start position	2.5
6)	‘나eliyeom’ thickness uniformity	2.75	47)	difference between character high and low point	4
Inclination			48)	talnemo feature	2.75
7)	vertical stem tilt	3.25	Serif		
8)	‘ㄱ ppichim’ shape	3.5	49)	curvature of consonant serif	3
10)	horizontal stem tilt	2.75	50)	slope of consonant serif	3
11)	‘ㄷ ppichim’ inclination	3.75	51)	existence of vertical stem serif	4.5
12)	‘ㅎ,ㅈ’ start stem inclination	3	53)	slope of double vertical stem	2.5
Consonant, Vowel			54)	curvature of single vertical stem	2.5
17)	horizontal and vertical stem thickness ratio	2.25	55)	existence of ‘ㅇ,ㅎ’ dolgi(start serif)	4.75
19)	‘ㄱ,ㅈ,ㅊ’ left-right symmetry	3.25	56)	‘ㅇ’ open or not	4
20)	vertical stem and double stem ratio	1.75	Rolling, Contact		
21)	‘ㅌ, ㅍ, ㅋ, ㆁ, ㆁ’ upper-lower symmetry	2	57)	‘ㄱ,ㅈ’ fork position	4.25
22)	double stem height difference	1.75	59)	existence of connected stem	3.5
23)	‘ㄱ,ㅍ,ㅇ,ㅈ,ㅎ’ left-right symmetry	2.75	60)	degree of flexagon ‘ㄴ,ㄷ,ㄹ,ㅁ,ㅂ’	4
24)	‘ㅌ, ㅋ’ side stem symmetry	2.25	61)	‘ㄴ,ㄷ,ㄹ,ㅁ,ㅂ’ connection point type	4
25)	‘ㄴ, ㄹ, ㅁ’ upper-lower symmetry	2	62)	existence of contact of double consonants	3
26)	‘ㄱ,ㄴ,ㄷ,ㄹ,ㅁ,ㅂ’ left-right symmetry	1.75	63)	contact with consonant and vowel	3
27)	top, middle, bottom alignment	2.25	64)	‘ㅍ’ internal opening (closure is common)	3.25
28)	choseong jungseong start position	1.5	65)	contact with horizontal and vertical stems (contact is common)	3
Space			Closing		
32)	choseong jungseong jongseong width ratio	2	66)	existence of stem closing (not right angle)	3.5
33)	choseong bounding-box ratio	2.75	Contour		
35)	‘ㄴ’ horizontal stem between-space ratio	2.25	71)	choseong jungseong width ratio	3.5

폰트의 모음 ‘ㅌ’의 가지 부분을 기준으로 굵기의 굵고 얇은 정도를 3단계(0-2-4)로 분류하였다. 네 번째는 돌출 획 요소에 위치한 맺음과 돌출이다. ‘ㄱ’과 ‘ㅂ’의 왼쪽 아래 부분을 기준으로 맺음 혹은 돌출

모양을 3단계(0-2-4)로 분류하였다. 다섯 번째는 꺾임 획 요소에 위치한 굴림과 꺾임이다. ‘ㄱ’과 ‘ㅂ’의 오른쪽 아래 부분을 기준으로 폰트의 등글거나 각진 정도를 4단계(0-1-2-4)로 분류하였다. 여섯 번째는

Stroke Element	Sample Image	Property Class Value
Sang-Tu		Shape of Dol-Gi 아 ... 아 0 ... 4
Bu-Li		Shape of vertical stem 아 ... 아 0 ... 4
Ga-Ji		Thickness of stem 아 ... 아 0 ... 4
Dol-Chul		Conclusion and Protrusion 마 ... 마 0 ... 4
Kkeokk-Im		Curviness and Angle 마 ... 마 0 ... 4
Ppi-Chim		Sharpness of Ppi-Chim 가 ... 가 0 ... 4
Kkog-Ji-Jeom		Shape of Kkog-Ji-Jeom 하 ... 하 0 ... 4

Fig. 8. Representative properties and measurement criteria.

빼침 획 요소에 위치한 빼침의 뾰족함이다. ‘ㄱ’의 빼침 획 요소의 모양의 날카로움 정도를 3단계(0-2-4)로 분류하였다. 마지막 일곱 번째는 꼭짓점 획 요소에 위치한 꼭짓점의 모양이다. ‘ㅎ’과 ‘ㅈ’의 꼭짓점 모양을 기준으로 4단계(0-1-3-4)로 분류하였다.

3.4 대표 폰트 선정

앞서 선정된 7개의 핵심 속성과 개별 획 요소 기반 유사성 계산 모델을 통한 영향력 분석 실험을 진행하기 위하여 대표 폰트를 선정하였다. 이 때 대표 폰트를 골고루 선정하기 위하여 선행 연구[9]에서 제시한 한국정보통신기술협회의 한글 모양 분류체계를 이용하였다. 그리고 분류 체계의 모든 항목에 대하여 무료로 접근이 가능한 폰트 중 각 항목을 대표할 수 있는 폰트를 선정하였다. 그에 따라 선정된 13개의

대표 폰트는 다음 Fig. 9와 같다.

3.5 유사성 측정 모델 설계

폰트 내 속성 파라미터값을 기반으로 폰트 간 유사성을 비교하기 위해서는, 각 폰트의 특징을 값으로 수치화하여 비교할 수 있어야 한다. 본 논문에서는 폰트별로 동일 속성 간 값의 차이를 유클리디언 거리 계산 방법을 사용하여 계산한 뒤, 이를 모든 속성에 대해 반복하여 그 총합을 폰트 간 특징 거리(property distance)로써 저장한다. 이 특징 거리가 짧을수록 폰트 간 유사성(font similarity)은 높게, 특징 거리가 멀수록 폰트 간 유사성은 낮게 평가할 수 있도록 개별 획 요소 기반 유사성 계산 모델을 설계하였다.

여러 폰트 간 유사성 비교 시 기준의 역할을 수행할 가상의 타겟 폰트를 t , 타겟 폰트와 비교하고자 하는 폰트를 비교 폰트 a 라고 가정한다. 이때 폰트 t 와 폰트 a 의 특징 거리를 구하는 공식은 다음 식 (1)과 같다.

$$prop_dist(t,a) = d_a^t = \sum_{i=k}^n \{|P_i(t) - P_i(a)| \times w_i\} \quad (1)$$

$$fnt_sim(t,a) = \frac{1}{prop_dist(t,a)} \quad (2)$$

비교하고자 하는 폰트 속성은 i 로 지칭하므로, $P_i(t)$ 는 타겟 폰트 t 의 속성 i 의 속성값이다. 같은 방식으로 $P_i(a)$ 는 비교 폰트 a 의 속성 i 에 해당하는 속성값이다. 이 둘의 값의 차이를 유클리디언 거리 계산 방법을 사용하여 구한 후, 해당 속성 i 의 중요도 가중치 w_i 을 곱한다. 중요도 가중치는 해당 속성이 폰트 특징 및 유사성 관별에 어떠한 비중으로 적용될 것인지에 대한 정보를 담고 있다. 따라서 개별 속성을 고려할 때는 가중치 값을 1, 고려하지 않을 때는 0으로 적용한다. 이 과정을 모든 핵심 속성 k 부터 n 까지 반복한 후 모든 값을 합산하여 나온 d_a^t 이 타겟 폰트 t 와 비교 폰트 a 간의 특징 거리(property distance)이다. 그리고 이 특징 거리를 역수로 변환한 것이 폰트 간 유사성 값(font similarity)이다. 이 유사성 값 클수록 타겟 폰트 t 와 비교 폰트 a 가 유사하고, 작을수록 유사하지 않다고 판단한다. 이는 식 (2)와 같다.

$$prediction = \min(d_a^t, \dots, d_a^t) \quad (3)$$

Large Category	Middle Category	Subdivision	Pont Name	Pangram
인체글씨 (Print Font)	부리글씨 (Serif)	경자체 (Print Style)	바탕체 (Batangche)	푸른 하늘과 빨간 태양의 따뜻한 햇볕 아래 붉은 귀뚜라미 보물 잔치
			HY 권명조 (HY Gyeon-Myeongjo)	푸른 하늘과 빨간 태양의 따뜻한 햇볕 아래 붉은 귀뚜라미 보물 잔치
			HY 신명조 (HY Sin-Myeongjo)	푸른 하늘과 빨간 태양의 따뜻한 햇볕 아래 붉은 귀뚜라미 보물 잔치
		모난부리체 (Angled Serif)	나눔 명조 (Nanum Myeongjo)	푸른 하늘과 빨간 태양의 따뜻한 햇볕 아래 붉은 귀뚜라미 보물 잔치
		둥근부리체 (Round Serif)	산돌 이야기 (Sandol Iyagi)	푸른 하늘과 빨간 태양의 따뜻한 햇볕 아래 붉은 귀뚜라미 보물 잔치
		응용부리체 (Applied Serif)	산돌 독수리 (Sandol Dongsuri)	푸른 하늘과 빨간 태양의 따뜻한 햇볕 아래 붉은 귀뚜라미 보물 잔치
	민부리 글씨 (San-Serif)	모난민부리체 (Angled San-Serif)	맑은 고딕 (Malg-eun Goding)	푸른 하늘과 빨간 태양의 따뜻한 햇볕 아래 붉은 귀뚜라미 보물 잔치
			돋움체 (Dod-umche)	푸른 하늘과 빨간 태양의 따뜻한 햇볕 아래 붉은 귀뚜라미 보물 잔치
			나눔 스퀘어 (Nanum Square)	푸른 하늘과 빨간 태양의 따뜻한 햇볕 아래 붉은 귀뚜라미 보물 잔치
		둥근민부리체 (Round San-Serif)	타이포 도담체 (Typo Dodamche)	푸른 하늘과 빨간 태양의 따뜻한 햇볕 아래 붉은 귀뚜라미 보물 잔치
			굴림체 (Gullimche)	푸른 하늘과 빨간 태양의 따뜻한 햇볕 아래 붉은 귀뚜라미 보물 잔치
			빙그레체 (Bing-geuleche)	푸른 하늘과 빨간 태양의 따뜻한 햇볕 아래 붉은 귀뚜라미 보물 잔치
			유니엣체 (Hyumeon-yesche)	푸른 하늘과 빨간 태양의 따뜻한 햇볕 아래 붉은 귀뚜라미 보물 잔치
			응용민부리체 (Applied San-Serif)	유니엣체 (Hyumeon-yesche)

Fig. 9. Representative fonts for similarity measurement experiments.

비교하고자 하는 폰트가 a 부터 z 까지 있을 때, 타겟 폰트를 기준으로 모든 비교 폰트와 계산한 속성값 기반 특징 거리는 d_a 부터 d_z 로 정리할 수 있다. 그리고 위 특징 거리값 리스트 중 가장 작은 값을 가진 비교 폰트가 타겟 폰트와 가장 비슷하다고 판단할 수 있다. 이에 기반을 두어 가장 작은 거리값을 가진 비교 폰트를 *prediction*에 저장하도록 한다. 위 과정을 식으로 나타낸다면 이는 식 (3)와 같다.

4. 실험 평가 및 분석

4.1 실험 계획

앞서 선정한 핵심 속성과 폰트 유사성 계산 모델은 폰트 간 유사성을 개별 획 요소를 기반으로 계산하여 유사성을 측정 및 비교한다. 그렇다면 이 방법이 실제 사람의 통합 인지 기반 방법과 일치하는 결과를 내는지 유효성을 검사할 필요가 있다. 더불어 각 속성별로 유사성 판단에 미치는 영향력을 파악하기 위해 실험을 진행하였다. 실험은 사용자의 통합 인지 방법과 개별 획 요소 계산 모델 방법별로 대표 폰트 간 유사성 순위를 구하고, 두 순위의 상대적인 일치도를 코사인 유사도를 사용하여 비교 및 평가하

는 것으로 구성되어 있다.

구체적인 실험 순서는 다음과 같다.

- ① [통합 인지 기반 방법 : 방법 1] 대표 폰트 간 유사성 설문 진행
- ② [통합 인지 기반 방법 : 방법 1] 설문 기반 대표 폰트 간 유사성 순위 도출
- ③ [개별 획 요소 계산 방법 : 방법 2] 개별 획 요소 기반 계산 모델을 사용하여 핵심 속성별로 대표 폰트 간 유사성 순위 도출.
- ④ [방법 1과 방법 2 비교] 통합 인지 기반 순위와 개별 획 요소 기반 계산 모델 순위 간 일치하는 정도를 코사인 유사도로 계산
- ⑤ [검증] 분산 분석을 통한 핵심 속성별 영향력 비교

4.2 통합 인지 기반 평가

20대부터 40대 이상의 41명의 여성과 31명의 남성으로 이루어진 총 72명의 참여자들이 13개의 대표 폰트에 대하여 유사성 판단 설문을 진행하였다. 유사성 판단 설문의 형식은 대표 폰트의 샘플 문장 이미지를 2개씩 묶어서 비교하고, 유사성 점수 평가에는

5점 리커트 척도를 사용하였다. 이때 보다 정확하고 공정한 평가를 위하여 설문 응답자에게는 대표 폰트의 샘플 문장 이미지 외에는 폰트명과 같은 폰트 관련 정보는 제공하지 않았다. 대표 폰트의 샘플 문장으로는 ‘훈민정음 스물여덟자’를 사용하였다.

$$Score(t,i) = \frac{\sum personalScore(t,i)}{num\ of\ appraisers(=72)} \quad (4)$$

$(1 < t, i < 13, t \neq i)$

$$ScoreList(t) = \{Score(t,i)\} \quad (5)$$

통합 인지 기반 방법에 의한 유사성 순위를 매기기 위하여 점수를 취합하였다. 13개의 대표 폰트 중 임의의 타겟 폰트 t 를 지정하고, 그 외의 12개의 대표 폰트를 비교 폰트 i 로 지정한다. 설문 참여자들이 폰트 t 와 i 에 대하여 평가한 유사성 점수를 합산하여 설문 참여자들의 수로 나누어서 평균을 낸 점수가 타겟 폰트 t 와 비교 폰트 i 간 통합 인지 기반의 유사성 점수가 된다. 이는 식 (4)와 같다. 이 과정을 12개의 모든 비교 폰트에 대하여 진행하여 생성된 리스트는 타겟 폰트 t 에 대한 점수 리스트가 된다. 위 과정은 수식 (5)와 같다.

$$\begin{aligned} ScoreRankMatrix &= \{ScoreRank(t) \mid 1 < t < 13\} \\ &= rank(\{ScoreList(t)\} \mid 1 < t < 13) \end{aligned} \quad (6)$$

앞서 임의의 타겟 폰트 t 에 대하여 도출한 점수 리스트를 대상으로 순위를 계산할 수 있다. 이때 같은 유사성 점수를 받은 경우는 같은 등수로 계산한다. 위 과정을 13개의 모든 대표 폰트에 대하여 진행하였을 때 통합 인지 기반 방법에 따른 유사성 순위 매트릭스를 만들 수 있다. 위 과정은 수식 (6)과 같고, 그 결과 매트릭스는 Fig. 10과 같다.

4.3 통합 인지 기반 평가와 개별 획 요소 기반 계산

모델의 비교

계산 모델로 도출한 유사성 순위와 통합 인지 기반 평가로 도출한 순위가 일치할 때 각 경우에서 고려한 획 요소가 중요하다고 판단할 수 있다. 이를 위해서는 두 순위 리스트의 일치 및 유사 정도를 수치화하여 파악하여 계산할 수 있어야 한다. 이 때 절대적인 값이 아닌 상대적인 유사성 순위가 서로 일치하는지 검증이 가능해야 한다. 따라서 상대적인 유사성 순위를 비교하기에 용이한 [19,20] 코사인 유사도 기법을 사용하였다.

이때 앞선 과정에서 구한 유사성 순위 리스트는 동일한 타겟 폰트를 기준으로 한 그 외의 비교 폰트 유사성 순위 값 리스트와 같은 형식으로 이루어져 있다. 즉, 각 비교 폰트들은 유사성 다차원 상에 있는 벡터의 차원인 축과 같은 역할을 하고 있다. 그에 따른 순위는 각 축에 대한 값의 역할을 한다. 이때 절대적인 값이 아닌 상대적인 유사성 순위가 서로 일치하는지 검증하기 때문에, 고차원에서 각도를 기반으로 상대적인 유사성 순위를 비교하기에 용이한 코사인 유사도 기법을 사용하기에 적합하다고 판단할 수 있다. 모든 유사성 순위 값은 양수이며, 각 벡터의 원소 수는 같기 때문에 코사인 유사도 기법을 사용하기 위한 전제 조건도 만족함을 확인할 수 있다.

$$\begin{aligned} \overrightarrow{ScRank} &= (f_{sr1}, f_{sr2}, f_{sr3}, \dots, f_{srn}) \\ \overrightarrow{DisRank} &= (f_{dr1}, f_{dr2}, f_{dr3}, \dots, f_{drn}) \end{aligned} \quad (7)$$

가상의 타겟 폰트 t 에 대한 유사성 순위 리스트를 하나의 벡터라고 가정한다면, 위 식 (7)과 같이 나타낼 수 있다. 사람의 통합 인지 기반 평가로 도출한 유사성 순위 벡터는 $ScoreRank$, 줄여서 $ScRank$ 라고 지칭하며, 타겟 폰트 t 에 대한 비교 폰트들의 순위 리스트인 f_{sr1} 부터 f_{srn} 까지의 목록으로 구성되어 있

	바탕체	HY견명조	HY신명조	나눔명조	산돌이야기	산돌독수리	맑은고딕	돋움체	나눔스퀘어	타이포도담체	굴림체	빙그레체	휴먼엠펜체
바탕체		2	1	1	5	6	6	8	7	9	6	5	7
HY견명조	3		2	2	4	5	10	9	10	8	8	7	6
HY신명조	1	1		1	1	5	8	7	6	10	7	6	7
나눔명조	2	3	2		2	5	5	4	6	7	4	4	6
산돌이야기	4	4	3	3		1	4	8	6	3	7	3	4
산돌독수리	9	6	7	8	6		9	10	8	6	7	5	3
맑은고딕	5	6	6	4	4	4		1	1	4	3	1	2
돋움체	6	5	4	2	7	6	2		2	4	1	4	4
나눔스퀘어	6	8	5	5	6	2	1	1		2	2	1	4
타이포도담체	9	6	9	7	3	1	4	6	4		4	1	1
굴림체	7	7	7	5	9	5	3	2	3	6		2	3
빙그레체	8	8	8	6	7	5	5	3	5	5	4		5
휴먼엠펜체	10	8	10	9	8	3	7	5	9	1	5	6	

Fig. 10. Score ranking matrix with integrated cognitive evaluation.

다. 같은 방식으로 개별 획 요소 기반 계산 모델로 도출한 유사성 순위 벡터는 *DistanceRank*, 줄여서 *DisRank*라고 지칭하며 f_{dri} 부터 f_{drn} 까지의 목록으로 구성되어 있다. 벡터 *ScRank*와 벡터 *DisRank*와 사이 끼인 각도를 세타(θ)라고 할 때, 코사인 세타 값으로 코사인 유사도 값 측정 및 비교가 가능하다. 이는 두 벡터의 내적 값을 벡터의 크기 값으로 나누면 계산이 가능하다. 구체적인 내용은 아래의 식 (8)과 같다.

$$\begin{aligned} \text{similarity}(\overrightarrow{ScRank}, \overrightarrow{DisRank}) \\ = \cos(\theta) &= \frac{SR \cdot DR}{\|SR\| \|DR\|} \\ &= \frac{\sum_{i=1}^n f_{sri} \times f_{dri}}{\sqrt{\sum_{i=1}^n (f_{sri})^2} \times \sqrt{\sum_{i=1}^n (f_{dri})^2}} \end{aligned} \quad (8)$$

모든 대표 폰트에 대하여 통합 인지 기반 유사성 순위와 개별 획 요소 계산 모델 기반 순위에 대하여 앞에서 소개한 코사인 유사도를 도출해내었을 때, 결과는 다음 그림 Fig. 11과 같다. 각 열에서는 개별 획 요소 기반 계산 모델에서 유사성 판단을 위해 어떤 획 요소의 속성을 고려한 것인지 보여준다. 각 행에서는 유사성 순위를 계산할 때 기준이 된 타겟 폰트가 무엇인지 보여준다.

4.4 글꼴 속성 영향력 분석

우선 대표 폰트의 부리(Serif), 민 부리(San-Serif) 계열별로 코사인 유사도 결과에서 어떤 양상을 보이는지 파악하고자 하였다. 부리(Serif) 글씨 계열로는

‘바탕체’, ‘HY견명조’, ‘HY신명조’, ‘나눔명조’, ‘산돌 이야기’, ‘산돌독수리’를 포함하고, 민 부리(San-Serif) 글씨 계열로는 ‘맑은고딕’, ‘돋움체’, ‘나눔스퀘어’, ‘타이포도담체’, ‘굴림체’, ‘빙그레체’, ‘휴먼엣체’를 포함한다. 부리 계열의 코사인 유사도 결과 평균 값과 민부리 계열의 코사인 유사도 결과 평균 값을 사용하여 그래프를 작성하였을 때 Fig. 11의 그래프와 같다. 7개의 개별 획 요소 중 꼭지점 획 요소를 제외한 모든 6개의 획 요소에서는 부리 계열과 민부리 계열 간 큰 차이 없이 비슷한 코사인 유사도 값을 보이는 것을 알 수 있다. 하지만 꼭지점 획 요소에는 부리 계열과 민부리 계열 간 코사인 유사도 값이 크게 차이 나는 것을 확인할 수 있다.

속성에 따라 코사인 유사도 계산 결과가 유의미한 차이가 있는지 확인하기 위해서 속성에 따라 나누어진 열을 기준으로 반복 없는 이원배치 분산분석을 진행하였다. 속성에 따라 유사성 판단에 미치는 영향력이 같다는 가설을 세웠을 때, 부리 계열에서는 p-value가 0.574913으로 확인되었다. p-value로는 가설이 맞을 확률을 확인할 수 있는데, 부리 계열의 p-value는 유의수준 0.05를 넘는 값으로 가설을 채택할 수 있다. 따라서 속성별로 미치는 영향이 통계적으로 차이가 없음을 알 수 있다. 하지만 민부리 계열에서는 p-value가 0.000437로 유의수준 0.05를 넘지 못해 가설을 기각하였다. 그에 따라 민부리 계열에서는 속성별로 미치는 영향이 통계적으로 다를 수 있었다. 즉 민부리 계열에서는 꼭지점 획 요소가 한 글 폰트 간 유사성을 판단하기에 다른 획 요소에 비

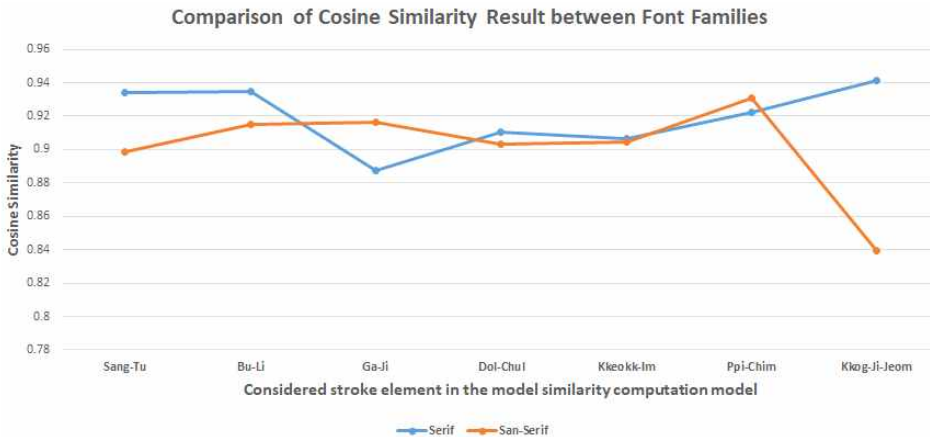


Fig. 11. Comparison of cosine similarity values of stroke elements by font family.

Table 2. Cosine similarity result of calculation model based on stroke elements

Considered Property		Cosine Similarity
Individual Property	Sang-Tu	0.91475332
	Bu-Li	0.92427478
	Ga-Ji	0.902886315
	Dol-Chul	0.906559427
	Kkeokk-Im	0.905496329
	Ppi-Chim	0.927026995
	Kkog-Ji-Jeom	0.886615287
All Property (together)		0.938591483

해서 적절하지 않음을 알 수 있다.

부리 계열과 민부리 계열을 구분하지 않고, 모든 폰트에 대하여 개별 획 요소별 코사인 유사도 결과를 속성 열을 기준으로 반복없는 이원배치 분산분석을 진행했을 때는 p-value를 0.346975로 확인할 수 있었다. 이는 유의수준을 0.05를 넘는 값으로, 그에 따라 귀무가설을 채택하였다. 따라서 계열 간 구분 없이 전체 폰트에 대한 유사성 판단 시에는 속성별로 미치는 영향은 통계적으로 유의미한 차이가 없음을 알 수 있었다. 각 속성 간 영향력에 차이가 없으므로, 모든 속성을 고려하였을 때의 결과를 다시 계산하여 기존 개별 획 요소를 고려했을 때의 결과와 비교하였다.

모든 속성을 고려했을 때의 코사인 유사도 결과는 0.938591483이다. 이는 기존 개별 획 요소만을 고려했을 때 가장 높은 코사인 유사도 결과 값을 도출한 삐침 획 요소의 결과값인 0.927026995 보다도 약 0.0115 정도 높은 값이다. 따라서 한글 폰트 간 유사성 판단 시에는 상투, 부리, 가지, 돌출, 꺾임, 삐침, 꼭지점 7개의 획 요소를 모두 고려하는 것이 개별적으로 고려하는 것 보다 좋은 결과를 낼 수 있음을 알 수 있다.

이에 더해 속성 조합에 따른 코사인 유사도 결과도 함께 비교해보았다. 핵심 속성 7개로 생성 가능한 조합을 모두 고려하여, 각 경우에 대한 코사인 유사도 값을 계산하였다. 그리고 고려한 속성 개수에 따라 코사인 유사도 값을 평균을 내었을 때 Fig. 12와 같은 그래프를 그릴 수 있다. 속성을 개별적으로 고려할 때보다는 속성 2개를 조합하여 고려할 때가 좋고, 2개를 조합할 때보다는 3개를 조합하는 것이 더 좋은 결과를 내었다. 즉 속성을 많이 조합할수록 코

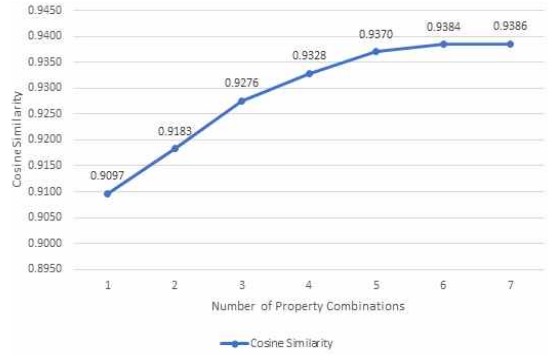


Fig. 12. Cosine similarity by Number of Property Combinations.

사인 유사도 값이 증가하여 그래프가 상승선을 그리는 것을 Fig. 12에서 확인할 수 있다. 따라서 보편적인 폰트들에 대한 유사성 평가 시에는 많은 수의 속성을 조합할수록 더욱 좋은 유사성 판단이 가능하며, 그에 따라 7개의 핵심 속성을 모두 고려할 때 가장 정확한 유사성 판단이 가능함을 알 수 있다.

5. 결 론

본 논문에서는 획 요소별로 한글 폰트 간 유사성 판단에 미치는 영향력을 파악하고자 하였다. 이를 위해 한글의 획 모양 요소에 기반을 두어 7개의 핵심 글꼴 속성을 선정하고 속성값 측정 방법을 정의하였다. 이를 기반으로 개별 획 요소 기반 유사성 계산 모델을 구축하였고, 본 모델의 검증 실험을 위해 13개의 대표 폰트를 선정하고 대표 폰트 간 유사성 순위를 계산하였다.

이후 사용자들에게 대표 폰트에 대한 유사성 설문을 진행하여 통합 인지 기반 유사성 순위를 도출해내었다. 그리고 코사인 유사도를 사용하여 통합 인지 기반 유사성 순위 결과와 계산 모델을 통한 순위의 상대적인 순위 일치 정도를 비교하였다. 최종적으로 분산 분석을 진행하여 획 요소 속성별로 미치는 영향력이 크게 다르지 않으며, 다수의 속성을 고려하는 것이 소수의 속성을 고려하거나, 속성을 개별적으로 고려하는 것 보다 좋은 결과를 낼 수 있음을 알 수 있었다.

본 연구에서는 속성을 단순하게 고려하거나 고려하지 않는 방식으로 유사성을 판단하기 때문에, 사용자의 평가와 같이 여러 속성의 중요도 가중치를 각각

실수로써 계산하는 것이 불가능한 상태이다. 만약 통합 인지 기반 결과와 가장 유사한 결과를 낼 수 있는 실수 가중치 값을 구하고, 각 속성을 선택하는 것에 대한 의미를 파악할 수 있게 된다면 유사성 판단 결과를 이전보다 향상시킬 수 있을 것으로 보인다. 더 나아가 획 요소와 함께 한글 구조를 이루는 골격과 공간 정보도 고려하여 유사성을 계산할 수 있게 된다면 더 정확한 유사성 판단이 가능해질 것이다.

향후 한글 획 요소 검출 및 속성 값 측정까지 자동화할 수 있게 된다면, 한글 폰트 간 유사성 비교의 자동화는 물론 한글 폰트의 정보 체계 전체를 자동화할 수 있게 될 것이다. 이는 곧 한글 폰트 연구의 활성화로 이어질 것이며, 더 나아가 타 언어의 폰트 간 정보 비교까지 용이하게 할 것으로 보인다. 최종적으로는 전 세계에서 함께 접근하는 웹페이지에서의 서로 다른 언어 간 유사 폰트 변환 자동화까지 가능해질 것이다.

REFERENCE

- [1] Korea National Hangul Museum, *Survey and Prospect of Font Industry in 2015*, Korea National Hangul Museum Research Report, 2015.
- [2] Y. Zhu, T. Tan, and Y. Wang, "Font recognition based on global texture analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 10, pp. 1192-1200, 2001.
- [3] X. Ding, L. Chen, and T. Wu, "Character Independent Font Recognition on a Single Chinese Character," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 2, pp. 195-204, 2007.
- [4] Z. Wang, J. Yang, H. Jin, E. Shechtman, A. Agarwala, J. Brandt et al., "Deep Font: Identify Your Font from An Image," *Proceeding of ACM International Conference on Multimedia*, pp. 451-459, 2015.
- [5] S. Jiang, Z. Wang, A. Hertzmann, H. Jin, and Yun Fu, "Visual Font Pairing," *arXiv Preprint arXiv:1811.08015*, 2018.
- [6] T. Chen, Z. Wang, N. Xu, and H. Jin, "Large-Scale Tag-based Font Retrieval With Generative Feature Learning," *Proceeding of IEEE/CVF International Conference on Computer Vision*, pp. 9116-9125, 2019.
- [7] KISDI, *In the Era of the 4th Industrial Revolution, Media Content Survival Strategy*, KISDI Premium Report 17-14, 2017.
- [8] H. Kim and S. Lim, "Shape Property Study of Hangul Font for Font Classification," *Journal of Korea Multimedia Society*, Vol. 20, No. 9, pp. 1584-1595, 2017.
- [9] H. Kim, *A Study on the Shape and Emotion Feature Analysis Algorithms for Hangul Font Recommendation*, Doctor's Thesis of Sookmyung Women's University, 2018.
- [10] PANOSE: An Ideal Typeface Matching System for the Web(1996), <https://www.w3.org/Printing/stevahn.html> (accessed August 20, 2020).
- [11] MyFonts(1999), <http://www.myfonts.com/> (accessed August 20, 2020).
- [12] FontPair(2010), <https://fontpair.co/> (accessed August 20, 2020).
- [13] D. Park, J. Jeon, Y. Ji, and S. Lim, "Design of Structure Map for Modeling Hierarchical Structure Information of Hangul," *Proceeding of the Fall Conference of the Korea Multimedia Society*, pp. 578-579, 2019.
- [14] T. Chen, Z. Wang, N. Xu, H. Jin, J. Luo, "Large-scale Tag-based Font Retrieval with Generative Feature Learning," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 9116-9125, 2019.
- [15] S. Shin, D. Kim, H. Park, B. Kang, and K. Sohn, "Finding Compact Class Sets for Korean Font Image Classification," *Proceeding of 2017 4th International Association of Pattern Recognition Asian Conference on Pattern Recognition*, Nanjing, pp. 226-231, 2017.
- [16] J. Jeon, J. Yang, and S. Lim, "Implementation of Automatic Hangul Handwriting Production

Service Using Deep Learning,” *Proceeding of The Korean Human Computer Interaction Society Conference*, pp. 1155-1157, 2019.

- [17] J. Jeon, Y. Ji, D. Park, and S. Lim, “Evaluation of Criteria for Mapping Characters Using an Automated Hangeul Font Generation System based on Deep Learning,” *Journal of Korea Multimedia Society*, Vol. 23, No. 7, pp. 850-861, 2020.
- [18] J. Jeon, D. Park, S. Lim, Y. Ji, and S. Lim, “Automatic Extraction of Hangeul Stroke Element Using Faster R-CNN for Font Similarity,” *Journal of Korea Multimedia Society*, Vol. 23, No. 8, pp. 953-964, 2020.
- [19] M. Kim and S. Lee, “Measures of Abnormal User Activities in Online Comments Based on Cosine Similarity,” *Journal of the Society for Information Security*, Vol. 24, No. 2, pp. 335-343, 2014.
- [20] C. Yoo, “Estimation of Cosine Similarity Change in Singular Value Decomposition,” *Journal of the Korean Society of Information Technology*, Vol. 18, No. 3, pp. 31-37, 2020.



박 동 연

2017년~현재 숙명여자대학교
ICT융합공학부
IT공학전공

관심분야: UI/UX, 기계학습, 딥러닝, 멀티미디어(폰트), 데이터분석 등



전 자 연

2019년 숙명여자대학교 ICT융합공학부 IT공학전공(학사)
2019년~현재 숙명여자대학교 IT공학과 학석사연계과정
2020년~현재 소설감성 창업

관심분야: 기계학습, 딥러닝, 데이터분석, 멀티미디어(폰트, 키네틱 타이포그래피), UI/UX 등



임 서 영

2017년~현재 숙명여자대학교 ICT융합공학부 IT공학전공

관심분야: 멀티미디어(폰트), 딥러닝



임 순 범

1982년 서울대학교 계산통계학과(학사)
1983년 한국과학기술원 전산학과(석사)
1992년 한국과학기술원 전산학과(박사)

1989년~1992년 (주)휴먼컴퓨터 창업 (연구소장)
1992년~1997년 (주)삼보컴퓨터 프린터개발부 부장
1997년~2001년 건국대학교 컴퓨터과학과 교수
2001년~현재 숙명여자대학교 IT공학과 교수
2015년~현재 숙명여자대학교 ICT융합연구소 연구소장
2006년 University of Colorado 방문교수
2014년~한국멀티미디어학회 회장, 명예회장
2012년~한국전자출판학회 회장, 명예회장
관심분야: 컴퓨터 그래픽스, 웹/모바일 멀티미디어 응용, 디지털 방송, 전자출판(폰트, 전자책, XML 문서), User Interface