

컴퓨터형 한글 서체 개발을 위한 자소 결합 알고리즘 연구

김윤식*, 엄정국**, 송만석*

*연세대학교 컴퓨터과학과

**나사렛대학교 전산정보학과

*wolfkim@december.yonsei.ac.kr mssong@december.yonsei.ac.kr

**jkeom@dove.nazarene.ac.kr

A study on the combination algorithm of Korean alphabet to develop the Hanguk fonts for computers

Yunsik Kim*, Jeongkook Eom**, Mansuk Song*

* Department of Computer Science, Yonsei University

** Department of Information and Computer, Korea Nazarene University

요약

컴퓨터 상에서 모든 한글 음절을 구현하고자 하면 현대한글 11,172음절의 완성형 코드나 조합형 코드를 사용해야 하는데 조합형의 경우 글자의 미려도가 떨어지는 문제점이 발생되므로 자소 별수를 늘려 그 문제점을 보완하려는 연구가 진행되어 왔다. 이는 메모리 및 코드처리상 비효율적인 요소가 많으므로 본 논문에서는 자소는 초·중·종성 각각 6벌씩만 제작한 후 자소의 어울림에 따라 자소의 이동과 변형으로 그 미려도를 추구할 수 있는 방안을 제시하고자 한다.

1 서론

컴퓨터가 사용자에게 많은 편의를 제공한다고 할지라도 한글을 모두 표현하지 못한다면 그 효용가치는 반감이 될 수밖에 없다. 한때 한글을 완성형 코드라는 제한을 두어 사용빈도가 높은 음절 2,500여자로 제한하여 초등학교 교과서에 나오는 글자¹⁾조차도 컴퓨터에서 사용할 수가 없었다 [1,2]. 조합형 코드 형태로 보완하여 모든 글자를 컴퓨터에서 구현할 수 있으나 조합형 코드를 사용할 경우 글자의 미려도가 너무 떨어지고 미려도를 높이려면 완성형코드로서 모든 음절을 정의해야 하는데 이 경우는 음절수²⁾가 너무 많아서 서체 제작이나 메모리 관리에 대한 낭비적인 요소가 많다. 본 논문에서는 완성형 형태로 제작되는 음절의 미려도 판단의 문제점을 제시하여 이

를 보완하여 조합형 형태에서 서체의 미려도를 높일 수 있는 컴퓨터형 한글서체 구현에 관한 알고리즘을 연구하고자 한다.

2 컴퓨터와 한글 코드

컴퓨터는 영어 사용을 기준으로 한 문자 자료 표현 방식으로 설계되었으므로 한글을 사용하기 위해서는 한글과 영어의 특성을 잘 분석하여 글자 코드 알고리즘 설계를 신중하게 해야 한다. 한글과 영어의 문자 구성의 특성은 <표1>과 같이 구별된다.

	한 글	영 문
글자 구성	자모의 어울림 단위로 구성	알파벳 한자 한자로 독립
의미 표현	어절 별로 의미 나타냄	알파벳이 모인 단어로 의미 나타냄
크기	모든 글자 같다	폭과 높이 서로 다름
중심선	모두 같다	다른 것도 있다
모양	자모의 모양이 틀릴 수 있다.	모든 알파벳이 같은 급수, 서체에서는 동일하다.

<표1> 한글과 영어 문자 특성 비교

이처럼 한글과 영어의 구별되는 특성을 분석해 보면 영어의 경우는 알파벳마다 각각 코드를 부여하였으므로 한글의 경우는 글자 구성의 최소단위인 음절을 기준으로 최종 코드를 부여해야 한다는 결론을 얻는데 영어 코드 부여방식과 비교한 현재 통용되는 한글코드 부여 개념은 <표2>와 같다.

- 1) ㄴ, ㄷ, ㅈ, ㅊ, ㅋ, ㆁ, ㅅ, ㅆ 등
- 2) 11,172음절로 받침 있는 글자 초성19×중성21×종성27에 해당하는 음절 10,773개와 받침 없는 글자 초성 19×중성21에 해당하는 399자

	한글	영어
코드부여기준	조합형 - 자모 완성형 - 음절	알파벳
크기	n바이트(2바이트 이상)	1바이트(7비트)
입출력 처리	입력코드, 처리코드, 출력코드, 별도처리	동일하게 처리

<표2> 한글과 영어의 코드 부여 방식

이 개념을 응용하여 컴퓨터 제조회사들은 제각기 다른 알고리즘으로 한글 구현을 하였는데 그 방법도 다양하고 정보관리와 데이터 호환성의 문제 등이 야기되어 그 통일적인 기준의 필요성을 느껴 국가 기관에서 1987년에 기준안을 마련하게 되었고 발생되는 문제점을 몇차례 보완하였다[2].

코드 이름	제정 년도	코드 형태	한글만의 특징	주관
KSC-5601-87	1987	2byte 완성형	최초의 국가 지정 표준코드 한글 2350자 지정	공업진흥청
KSC-5601-92	1992	2byte 조합형	KSC-5601-87에 추가 현대 한글 음절 11,172자 추가	공업진흥청
KSC-5700-95	1995	완성형 조합형	한글의 입출력을 컴퓨터 내부에서 처리할 수 있는 코드 체계로서 조합형 형식으로 사용 가능한 자모 240개와 완성형 음절 11,172개 모두로 구성된 국제 문자 부호계를 한글 표준 문자계로 제정 고시	공업진흥청

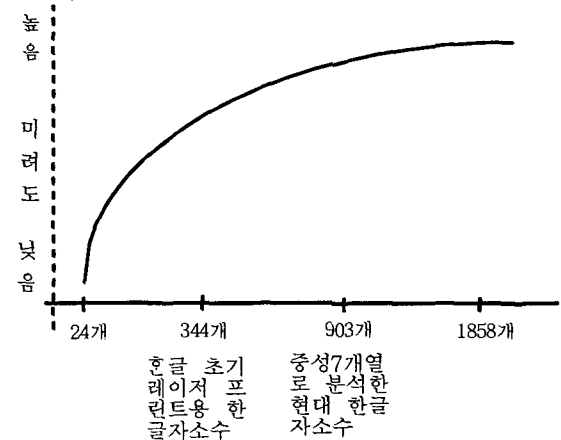
<표3> 한글 코드제정 변천

컴퓨터상에서 한글을 구현하는 최상의 방법은 모든 음절을 완성형 상태로 코드값을 부여하는 것인데 이는 코드를 부여할 영역이 커야 하고 서체 한벌을 제작하는데 비용이 너무 많이 소모되어 비효율적이므로, 한글 중에서 특히 자주 사용되는 음절을 선택하여 일부 음절만 컴퓨터에 사용할 수 있게 하여 한글 음절중 80%에 해당하는 글자를 사용할 수 없게 되었으며, 기계사용 관점으로 보면 한글을 컴퓨터에서 사용하는데 간편하게 하고자 글자수를 제한하는 것인데 이는 문화창달과 계승의 근간이 되는 언어의 다양성을 감소시키므로 적

은수의 코드 할당으로 모든 글자를 표현할 수 있는 방안으로 고안된 것이 조합형 코드 방식이다. 다음과 모음인 자소별로 7bit 코드를 부여하여 자소를 조합한다면 초성자음 19개 중성모음 21개 중성자음 27개의 코드값만 부여하여 11,762개의 모든 글자를 구현할 수 있으므로 입력코드 처리에 대한 효율의 향상을 보인다. 그러나 한글의 자소 모양은 초·중·중성의 어울림에 따라 그 모양새가 달라 지므로 하나의 자소에 하나의 코드만을 부여할 경우는 글자의 미려도가 저하될 수밖에 없다. 컴퓨터에서 모든 한글을 표현할 수 있어도 글자의 모양이 미려하지 못하다면 보완점을 찾아야 한다. 이를 위해 한글코드를 입력코드, 처리코드, 출력코드로 나누어 처리하는 방식을 KSC-5700에서 도입하였다. 입력과 처리 방식에서는 조합형코드를 사용하고 출력은 절충형 코드를 사용했다[1]. 즉, 초·중·중성을 순서대로 입력받아 내부 처리 알고리즘에 의해 하나의 음절로 조합한 후 이 음절에 해당하는 글자가 자주 사용하는 글자이면 완성형 코드에 해당하는 글자를 출력하고 그렇지 않으면 자소 조합형 글자를 출력한다. 그러나 이런 방법도 조합형 글자로 출력되는 글자는 미려도가 떨어 지므로 근본적인 해결점이 되지 않는 것이다.

3 자소코드와 미려도

한글 자소의 경우 'ㄱ'이라도 '가', '감', '깨' 등에서 서로 모양새가 다른데 음절의 미려도를 고려한다면 조합형일 경우 하나의 자소라도 초·중·중성의 어울리는 자소의 조합에 따라 각각 다른 코드값을 부여해야 한다. 리스곡선(Leeks Curve) 보면 자소의 갯수를 많이 구분할수록 미려도가 좋아진다[1,3,4]



<표4> 리스곡선

한글의 가독성과 변별력만 고려하면 24개의 자모로서 모든 현대 한글 음절을 표현할 수 있으나 최소 344개 정도의 자소를 조합해야 어느 정도 모양새를 갖춘 음절이 표현되며 903개의 자소를 조합하면 단행본에서 사용되는 음절정도의 미려도가 표현된다.

4 최소 자소 별수를 이용한 미려도 추구 방안

1858개의 자소를 사용할 경우는 현대 한글 음절에 포함되는 자소 33,117⁷⁾개를 사용했을 때[4]와 거의 같은 정도의 미려도를 추구할 수 있다는 것을 리스콕선을 보면 알 수 있다. 그러나 무작정 자소수를 늘리기 보다 834⁴⁾개의 자소로서 자소이동으로 미려도를 1858⁵⁾ 개 자소를 사용한 것에 가까운 미려도를 추구한 연구도 있었으나[3,4,5] 이 모든 표현 방법은 완성형 형태의 글자의 모양새를 기준으로 하였는데 이 기준 원도 역시 과학적이고 객관적인 미려도라기 보다는 몇몇 서체 디자인 전문가나 서예가의 주관적으로 관점에 특정인이 작성한 서체 원도에 의해 제작되었다[6,7].

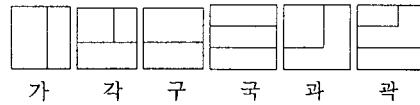
하나의 자소가 한 음절이 되는 일본 글자같은 음절단위의 서체제작이 타당할 것이나 자소의 어울림으로 하나의 음절을 형성할 경우에는 각각 자모에 다른 결합요소의 배치적인 미려도를 논리화해야하므로 이제부터라도 객관적이고 과학적인 방법으로서의 서체연구, 특히 컴퓨터에 의한 자모의 어울림과 미려도를 추구할 수 있는 방안이 연구되어야 한다. 한글의 서체는 현재의 인쇄술이 일본의 영향을 받아 자모가 어울리는 모아쓰기 글자의 특성이 배제되고 한 글자 한 글자의 미려도를 고려해 작성한 원도를 디지털화하여 폰트를 제작하는 방법을 답습했다. 즉, 한국의 서체연구가⁶⁾가 쓴 각 음절의 원도를 일본이 가져가서 디지털 폰트로 제작하여 모리사와, 사연 등의 일본회사에서 컴퓨터와 함께 국내로 역수출하였는데 이 방법의 서체 제작은 모아쓰기 형태의 서체 제작에는 비효율적인 면이 있다.

이에 반하여 1970년대 말부터 컴퓨터 보급으로 인해 모아쓰기 기준 서체의 필요성이 대두되어 자소 기준의 서체가 개발되었으나 미려도를 제대로 갖추었다는 것은 탈네모형태의 서체이고 네모형태의 서체는 컴퓨터 업계의 주관하에 진행되어 왔으나 미려도가 음절기준으로 개발된 서체에 비해 현저히 떨어졌다 [8]. 그러나 이러한 미려도 저하는 각각의 자소를 위치와 형태를 고정시킨 상태에서 단순히 자소수만 늘려 미려도를 추구하고자 하는데 문제점이 있어 자소의 변화와 위치 이동 개념의 도입이 필요하다.

각과 건에서 ‘ㄱ’을 굳이 틀리게 해야 아름답게 본다는 것은 폰트의 미려도보다는 예술적 감각과도 같은 개성일 수 있다. 본 논문에서는 한글의 특성을 잘 반영한 자소의 조합 알고리즘을 통한 한글 자소의 컴퓨터형 서체의 미려도를 추구할 수 있는 폰트제작에 대해 제시한다.

- 3) 중성이 없는 음절 399자의 798자소, 중성이 있는 음절 10,773자의 32,319자소
- 4) 1993년 김진하에 의해 제안된 자모
- 5) 1992년 오정급에 의해 제안된 글자꼴로 초성 52별, 중성 8별, 종성 26별로 이루어졌으며 중성을 13개열 나눈 점과 받침에 ‘ㄴ’이 올 경우를 다시 세분화했다.
- 6) 최정순, 장봉선 등

우리 한글은 크게 다음 6가지의 자모의 구성형태를 가지고 있다.



<표5> 한글 자소 결합 형태

물론 <표5>에서 처럼 모든 초·중·종성의 자소가 위치하는 영역이 분리된다는 것은 아니다. <표5>는 그 형태의 기본적인 모양새를 구분하기 위한 것이다. 즉, <표6>처럼 ‘각’ 형태의 글자에서 처럼 초·중·종성의 영역이 겹치면서 음절이 형성되는 것이다.



<표6> 초·중·종성자소의 영역겹침

본 논문에서는 미려도를 위한 자모를 제작하는 별수를 최소로 한 후 이 자모를 이동 및 변화를 주어 미려도를 추구하는 방안을 제시하고자 한다. 조합형으로 최대의 미려도를 추구한 논문[3]의 자모수와 본 논문의 자모수의 구성은 <표7>과 같다.

	1858자소	402자소
초성	52별×19=988	6별×19=114
중성	8별×21=168	6별×21=126
종성	26별×27=702	6별×27=162

<표 7> 자모 별수 비교

표에서 나타난 것과 같이 1858개 자소를 402개로 줄여서 사용하니 폰트제작 및 메모리 낭비 80%이상 줄일 수 있고 미려도가 떨어지지 않도록하기 위해서는 모음종류별 구분, 어울리는 빗침 상호방향, 자모사이의 여백의 분포 등을 변수로 하여 자모의 상하이동, 축소확대, 격입 각도의 변화를 이용한 미려도를 추구하고자 한다.

4.1 초성

받침없음, 단받침, 단받침중 ‘ㄴ’일 경우, 복받침의 5그룹으로 하고 모음을 13개⁷⁾ 그룹으로 묶어 52별수로 했다[3,4].

여기서는 모음 ‘ㅏ ㅑ’ 그룹과 ‘ㅓ ㅕ’ 그룹의 경우 같은 초성을 사용하더라도 초성과 중성 사이의 여백을 조정하고자 모음을 좌우로 이동한 후 ‘가’일 경우보다 ‘거’일 경우 ‘ㄱ’을 조금 축소하면 되겠다.

4.2 중성

7) (ㅏ ㅑ), (ㅓ ㅕ), (ㅗ ㅛ), (ㅜ ㅠ), (ㅡ ㅟ), (ㅣ ㅢ), (ㅝ ㅞ), (ㅝ ㅞ), (ㅟ ㅠ), (ㅢ ㅣ), (ㅣ ㅤ), (ㅥ ㅦ), (ㅧ ㅨ), (ㅩ ㅪ), (ㅫ ㅬ), (ㅭ ㅮ), (ㅯ ㅰ), (ㅱ ㅲ), (ㅳ ㅴ), (ㅵ ㅶ), (ㅷ ㅸ), (ㅹ ㅺ), (ㅻ ㅼ), (ㅽ ㅿ), (ㅿ ㅿ), (ㅿ ㅿ)

받침없음, 단자음 받침, 복자음, ㄴ받침 등 중성 4 그룹 단자음, 복자음 2그룹으로 8별을 제작해야 하나 '가'와 '까'의 예를 들면 'ㄱ'을 우측으로 이동하고 논문[6,7]에서 사용한 오른 결줄기(-)를 조금 줄이면 된다.

4.3 중성

단자음 초성, 복자음 초성 2그룹, 모음 13그룹으로 26별을 제작해야 한다.

'곡'과 '국'의 예를 들면 국의 경우 세로줄기를 좀 더 축소하고 아래로 이동하면 된다. 이상과 같이 하나의 자소로서 어울림에 따라 보정하는 방안에 대해 예를 들었다. 이러한 방법이 자소의 어울림에 따라 음절이 생성되는 글자의 특성을 그대로 컴퓨터에서 반영한 방안이다.

완성형 형태나 1858개의 자소결합의 글자라도 미려도는 한 음절 한 음절의 자체로는 아름답다고 하겠으나 전후 좌우의 글자와 어울리는 형태로는 어색한 모양을 보일 때가 많다. 자소의 변형 및 이동 알고리즘의 개발여부가 자소의 별수 개발보다 우리 한글의 미려도를 추구하기 위한 합리적인 방향이다.

또한 추후의 연구는 한 음절의 미려도 보다도 어절간의 글자 어울림을 고려한 자소를 이동 및 변형에 의한 미려도 추구 연구도 있어야 한다.

5. 결론

컴퓨터가 편리한 정보화 시대의 필수적인 기기라 할지라도 한글의 모든 음절을 구현할 수 없다면 한글 사용자에게는 아무 소용이 없다. 그렇다고 음절 각각을 하나하나 폰트로 제작하기란 너무 비효율적이므로 조합형 형태로서 최소의 자모 별수를 제작한 후 이 자소들을 자모가 어울리는 특징에 맞게 이동과 변형을 주어 이를 규칙화 하여 알고리즘을 개발하면 자소를 제작하는 별수를 줄일 수 있으면서 글자체의 미려도를 향상시킬 수 있어 매우 효과적이 될 수 있어 여러 서체를 제작하는데 비용과 시간이 절감되며 컴퓨터에서 우리의 아름다운 한글을 구현할 수 있을 것이다.

감사의 글

본 연구를 수행하는데 도와주신 계원예술대학 이기성 교수, 나사렛대학교 한경호박사, 연세대학교 한글정보처리연구실 양단희 연구원께 감사드립니다.

참고문헌

[1] 이기성, 전자출판 시스템중 CTS용 한글 음절 출력 방식에 관한 연구, 석사학위논문, 단국대, 1991

[2] 손애경, 전자출판에 있어서의 바람직한 한글 코드 설정에 관한 기초적 제언-초·중·고등학교 국어교과서의 음절 출현 분석을 중심으로-,

석사학위논문, 동국대, 1990

[3] 오정금, 자소조합에 의한 전자출판용 본문체 개발 및 미려도 연구, 석사학위논문, 동국대, 1992

[4] 김진하, 디지털 자소의 위치 이동에 의한 경제적인 CTS용 한글 글자꼴 구현 방식에 관한 연구, 석사학위논문, 동국대, 1993

[5] 이기성, 한글 출력코드의 릭스곡선에 관한 연구, 92 출판연구, 범우사

[6] 이기성, 전자출판-II, (주)장왕사, 1997

[7] 장은정, 한글 명조 활자체에 관한 연구, 석사학위논문, 서울여대, 1991

[8] 김진평, 한글 활자체의 조합형 설계 가능성 연구, 한글 및 한국어 정보처리 학술대회, pp. 293~300, 1992