

특허 인용에 영향을 미치는 요인 분석*

Analysis of Factors Influencing Patent Citations

유재복(Jae-Bok Yoo)**

정영미(Young-Mee Chung)***

초 록

최근 특허기술의 가치평가가 크게 강조되고 있으며, 그 평가의 수단으로 특허의 피인용횟수가 매우 유용한 척도 중의 하나로 받아들여지고 있다. 그에 따라 이 연구에서는 특허의 피인용횟수와 이에 영향을 미칠만한 형태적·기술적·개념적 요인의 17개 변수들 간의 상관관계를 미국특허를 대상으로 5개 주제분야에 걸쳐 분석하였다. 분석결과 특허의 피인용횟수와 일정 수준 이상의 상관관계, 즉 5% 이상의 설명력을 갖는 변수는 페이지 수, 청구항 수, 참고문헌 평균 피인용횟수, 기술분야 특허증감율, 서지결합도, 동시인용도 및 문헌간유사도 등 7개로 나타났다. 또한 이들 변수에 대한 분산분석 결과 7개 변수 모두 전반적으로 대부분의 주제분야 간에 있어서 평균값의 차이가 있는 것으로 나타났다.

ABSTRACT

Recently, the valuation of patented technology has been greatly emphasized, and patent citation has been accepted as a very useful index of this technology. In this study, we performed correlation analyses between the patent citation counts and 17 explanatory variables of morphological, technological, and conceptual factors with a test dataset of U.S. patents in five subject fields. Seven variables having 5% or more standardized variances(r^2) with patent citation counts were identified: number of pages, number of claims, reference-average-citation rate, patent increase/decrease rate, strength of bibliographic coupling, co-citation counts and document similarity. The result of the ANOVA test shows that the mean values of these variables vary among most subject fields.

키워드: 특허인용, 인용분석, 특허 피인용횟수, 특허인용 영향요인
patent citations, citation analysis, patent citation counts, factors influencing patent citation

* 이 논문은 박사학위 논문의 일부를 요약·정리한 것임.

** 한국원자력연구원 책임연구원(jbyoo@kaeri.re.kr) (제1저자)

*** 연세대학교 문헌정보학과 교수(ymchung@yonsei.ac.kr) (공동저자)

■ 논문접수일자: 2010년 2월 8일 ■ 최초심사일자: 2010년 2월 28일 ■ 게재확정일자: 2010년 3월 13일
■ 정보관리학회지, 27(1): 103-118, 2010. [DOI:10.3743/KOSIM.2010.27.1.103]

1. 서론

1.1 연구의 목적

지식기반 및 글로벌 경제로 대표되는 현대 사회에서 지식과 정보는 부의 수단에서 그 자체가 부의 원천이 되고 있으며, 토지나 자본 등 유형자산보다는 기술력, 브랜드, 디자인 등과 같은 무형의 지식재산이 국가나 기업의 경쟁력의 핵심요소로 등장하고 있다. 그 가운데 특허는 무형자산의 대표적인 지적 산물로 개인과 기업 및 국가의 기술수준과 혁신역량을 가늠하는 데 있어 중요한 객관적인 척도로 활용되고 있다. 그에 따라 각 기업이나 국가의 투자 및 정책 결정에는 물론 기술이전, 라이선싱, 제품 개발 등에 있어서 하나의 수단으로서 특허기술에 대한 가치평가가 크게 강조되고 있지만 특허기술의 가치를 객관적이고 정량적으로 측정하는 것은 무척 어려운 일이다.

특허기술의 가치를 평가하는 수단으로 여러 가지 정량적인 척도가 제안되어 있지만, 기존의 많은 실험적 연구결과 특허의 피인용횟수가 보다 중요한 척도로서 가장 광범위하게 받아들여지고 있으며(Dutta and Weiss 1997; Moorman and Slotegraaf 1999; Lin, Chen, and Wu 2007), 자주 인용된 특허는 보다 높은 기술적·경제적 가치를 창출하는 것으로 나타났다(Narin, Carpenter, and Woolf 1984; Trajtenberg 1990).

그러나 특허는 논문과 마찬가지로 발표, 즉 공개 또는 등록된 후 일정 기간이 경과되어야 활발하게 인용되는 특성으로 인해 최근에 출원되어 공개 또는 등록된 특허의 피인용횟수를 제대로 파악할 수 없다는 문제점이 있다. 즉 특

허가 후행 특허로부터 더 많은 인용을 받기 위해서는 일정 기간이 경과되어야 하는데, Hall 등(Hall, Jaffe, and Trajtenberg 2001)에 따르면 특허의 누적 피인용 비율이 50%에 이르기까지는 특허가 공개된 후 최소한 10년 이상 경과되어야 한다고 발표하였다. 결국 특허의 가치를 측정하는 중요한 척도인 피인용횟수는 이처럼 시간상의 제약으로 인해 즉각적인 활용이 어렵고, 특히 최근에 발표된 특허의 경우 사실상 피인용횟수를 정확하게 파악하는 것은 거의 불가능하다고 할 수 있을 것이다.

그에 따라 최근의 새로운 특허에 대한 피인용횟수를 예측할 수 있는 모형 개발의 일환으로 특허의 인용에 영향을 미치는 여러 가지 요인들에 대한 연구가 진행되고 있는데, 지금까지의 연구는 주로 프론트 페이지(front page)에 나타난 일부 데이터만을 대상으로 한 연구가 부분적으로 수행되었을 뿐 종합적인 분석은 이루어지지 않았다. 게다가 해당 주제분야의 기술발전적인 측면을 반영한 기술적 요인에 대해서는 거의 연구가 진행되지 않았으며, 주제 개념의 특성을 반영하는 개념적인 요인의 경우 전혀 연구가 이루어지지 않았다.

이 연구의 목적은 최적의 특허인용 예측모형을 개발하기 위한 기초적 자료를 제공하는 데 있다. 이를 위해 국가과학기술위원회에서 선정한 차세대 5대 핵심 기술분야를 대상으로 특허의 인용에 영향을 미칠만한 제반 변수들을 크게 형태적 요인, 기술적 요인, 개념적 요인 등 3가지 측면으로 나누어 종합적으로 분석하였다. 그 결과 어떠한 변수들이 특허의 피인용횟수와 상관관계가 있는지를 찾아내고 그러한 변수들이 주제분야별로 차이가 있는지를 검토하였다.

1.2 연구의 방법 및 범위

이 연구에서는 특허의 인용에 어떠한 요인들이 얼마만큼의 영향을 미치는가를 종합적으로 살펴보고자 2001년 12월 15일 과학기술기본법 제7조에 의거 국가과학기술위원회에서 선정한 미래유망 신기술 6T분야, 즉 BT(생명공학기술), CT(문화기술), ET(환경기술), IT(정보기술), NT(나노기술), ST(우주항공기술) 분야 중 CT분야를 제외한 5T분야를 대상으로 분석하였다. CT분야는 최근 인터넷의 활성화와 디지털 기술의 발전으로 급격히 부상하고 있는 기술로서 누적된 특허건수가 많지 않고 IT분야에 포함될 수 있기 때문에 분석대상에서 제외하였다. 주제분야별 분석대상 범위는 각 분야별로 등록된 특허건수의 형평성을 고려하여 그 범위를 적절하게 축소 또는 확대하였으며, 각 분야별 세부 주제분야는 <표 1>과 같다.

이 연구에서의 분석대상 특허는 미국특허로 한정하였다. 미국특허는 다른 나라의 특허와는 달리 다량의 폭넓은 인용정보를 기계가독형태로 제공하고 있고, 미국특허의 발명자 중 절반 가량이 미국인이 아닌 외국인이 차지할 만큼(노경란 2006) 세계 기술시장을 대표하는 특성을 가지고 있어서 세계 각국의 중요한 신기술은

거의 대부분 미국특허로 권리화하고 있기 때문이다(Narin, Hamilton, and Olivastro 1997; Iversen 2000).

이 연구에서는 특허의 피인용횟수에 영향을 미치는 변수들을 종합적으로 살펴보고자 형태적 요인, 기술적 요인, 개념적 요인 등 3가지 측면에서 17개의 변수를 토대로 연구가설을 설정하였다. 연구가설을 검증하기 위해 이 논문에서는 SPSS for windows 12.0 version 프로그램을 이용하여 종속변수인 특허의 피인용횟수와 17개의 독립변수들을 대상으로 상관관계분석을 실시하였다. 상관관계분석 결과, 특허의 피인용횟수에 일정 수준 이상의 상관관계가 있는 변수들을 대상으로 그들 변수들이 주제분야별로 평균의 차이가 있는지의 여부를 확인하고자 일원분산분석(one-way ANOVA)을 실시한 후 Scheffe 사후검증을 수행하였다.

1.3 선행연구

특허의 피인용횟수와 기술적·경제적 가치간에 상당한 상관관계가 있다는 것이 많은 연구를 통해 밝혀짐에 따라 특허의 인용에 영향을 미치는 요인이 무엇인지에 대한 연구가 최근 들어 진행되고 있다.

<표 1> 주제분야별 세부 주제분야

주제분야	세부 주제분야
BT(생명공학기술)	'미생물·효소; 미생물의 보존·유지·증식(C12N)' 분야
ET(환경기술)	전기소자 중 '화학에너지의 전기에너지 직접 변환 방법(H01M)' 분야
IT(정보기술)	'다중통신(H04J)' 분야
NT(나노기술)	'나노기술' 응용분야
ST(우주항공기술)	'항공기·비행·우주공학(B64)' 분야

Lanjouw와 Schankerman(2004)은 1975~1993년에 미국 제조회사에서 7개 기술분야에 걸쳐 출원된 10만여 건의 특허를 대상으로 특허의 질과 연구생산성과의 관계를 분석하는 과정에서 특허의 질을 평가하는 변수로 4개의 독립변수, 즉 청구항 수, 참고문헌 수, 인용문헌 수, 출원국가 수 등을 분석하였다. 그 결과, 특허의 질은 회사의 주식시장 가치와 긍정적인 상관관계가 있다고 밝히면서 특히 청구항 수가 7개 기술분야 가운데 6개 분야에서 가장 중요한 변수로 나타났다고 밝혔다.

Lee 등(Lee, Lee, and Song 2006)은 한국 전자통신연구원(ETRI)에서 미국 특허청에 출원하여 등록된 1,000여건의 특허를 대상으로 피인용횟수에 영향을 미치는 요인을 3가지 측면, 즉 연구팀 관련 변수(발명자 수, 자기인용 수, 단독-공동 출원여부), 발명 관련 변수(청구항 수, 출원국가 수, 비특허인용 수, 발명의 주제분야) 및 인용국가 관련 변수(미국특허에서의 인용, 일본특허에서의 인용, 기타국가에서의 인용) 측면에서 NB(negative binominal), ZAP(zero-altered poisson), ZANB(zero-altered negative binominal) 등 3가지 모형을 이용하여 분석하였다. 그 결과, 자기인용 수와 청구항 수, 전자분야에 속한 특허 및 미국특허에서의 인용 등이 ETRI 발명특허의 피인용횟수에 긍정적인 영향을 미치는 반면, 출원국가 수는 오히려 부정적인 영향을 미친다는 것을 밝혀냈다.

Lin 등(Lin, Chen, and Wu 2007)은 14개 바이오기술분야의 미국특허를 대상으로 5개의 독립변수에 대한 단순회귀분석을 통해 특허의 출원인 국적, 지리적 위치, 청구항 수 및 참고문

헌 수 등은 특허의 피인용횟수와 통계적으로 유의미한 상관관계에 있는 반면, 심사기간은 부분적으로 긍정적 또는 부정적인 상관관계가 있음을 밝혀냈다.

2. 연구설계

2.1 데이터 수집 및 전처리

이 연구에서의 각종 분석에 필요한 기본적인 데이터는 주로 (주)웹스의 특허검색시스템에서 제공하는 IPC(International Patent Classification, 이하 IPC) 탐색기능을 통해 검색한 후 '다운로드' 기능을 통해 수집하였다. 단, NT분야의 경우 별도의 IPC가 부여되어 있지 않아 키워드 검색을 통해 데이터를 확보하였다. 또한 (주)웹스에서 제공하고 있지 않으나 일부 변수를 산출할 때 필요한 1976년 이전의 미국 등록특허에 대한 각종 데이터는 미국특허청(United States Patent and Trademark Office)의 특허검색 사이트를 통해 수집하였다.

분석대상 특허의 등록년도는 주제분야마다 상이하다. 주제분야별로 특허의 피인용반감기(cited half-life)를 적용하여 분석대상 기간을 설정하였기 때문이다. 일반적으로 특허의 경우 해당 주제분야의 기술발전 속도가 빠를수록 피인용반감기가 짧고 느릴수록 피인용반감기가 길어지는 경향이 있다. 이에 이 논문에서는 피인용반감기에 대한 주제분야별 차이로 인한 형평성 문제를 없애고자, (주)웹스의 특허검색시스템에서 제공하는 미국등록특허 중 1976.1.1~2008.12.31까지의 특허를 대상으로 각 주제분

아별로 검색한 결과 수집된 인용정보를 토대로 분석대상 세부 주제분야에 대한 특허 피인용반감기를 산출·적용하여 분석대상 기간을 설정하였다. 여기에서의 피인용반감기는 JCR(Journal Citation Reports)에서 사용하는 아래의 피인용반감기 산출방법을 토대로 산출하였다.

$$\text{피인용반감기} = \text{누적인용률 50\% 이전 연도 수} + \frac{\text{누적인용률 50\%에서 50\% 직전 연도 누적인용율을 뺀 값}}{\text{50\% 직후 연도 누적인용률에서 50\% 직전 연도 누적인용률을 뺀 값}}$$

〈표 2〉는 주제분야별 특허 피인용반감기를 적용한 분석대상 특허의 등록년도 및 특허건수이다. 분석대상 특허건수는 BT분야 5,337건, ET분야 5,929건, IT분야 4,354건, NT분야 2,313건 그리고 ST분야 3,742건 등 총 21,675건이다.

2.2 변수 및 데이터 산출

이 연구에서 일부 변수의 데이터는 특허검색 시스템을 통해 다운로드한 값을 그대로 사용함

반면 대부분 변수의 데이터는 수작업 또는 별도로 설계한 프로그램을 통해 산출하였다. 〈표 3〉은 이 연구에서 필요한 각종 변수 및 데이터에 대한 산출방식을 정리한 것이다. 주요 변수와 데이터의 산출방법에 대한 상세한 설명은 필자의 학위논문에 제시되어 있다.

참고로, 기술적 요인의 변수 중 ‘기술분야 피인용반감기’, ‘기술분야 특허집단 크기’ 및 ‘기술분야 특허증감율’ 등 3개의 변수는 해당 변수 값을 산출할 때 IPC의 메인그룹(main group) 단위로 분석하였으며, 등록된 특허건수가 20건 이하인 기술분야는 분석대상에서 제외하였다. IPC의 메인그룹 아래 서브그룹(subgroup) 단위에서 분석할 경우 등록된 건수가 너무 적고, 등록된 특허건수가 20건 이하인 경우 해당 기술분야를 대표하기에는 다소 미흡하다고 판단했기 때문이다. 아울러, ‘기술분야 특허증감율’의 경우 각 기술분야별 세부 기술분야를 다시 특허의 등록시기에 따라 5개 구간으로 나누어 분석하였으며, 특정 구간에서 특허 등록건수가 0인 경우에는 증감율을 산출할 수 없는 관계로 분석대상에서 제외하였다.

또한 개념적 요인의 3가지 변수는 〈표 3〉의 데이터 산출방법에 따라 서지적 또는 개념적으

〈표 2〉 주제분야별 분석대상 특허 등록년도 및 특허건수

주제 분야	전체 특허건수 (1976~2008년)	반감기	분석대상 특허	
			등록년도	특허건수
BT	21,130건	12년(12.35년)	1976년~1996년	5,337건
ET	18,134건	14년(14.10년)	1976년~1994년	5,929건
IT	14,868건	13년(13.07년)	1976년~1995년	4,354건
NT	11,514건	11년(10.53년)	1976년~1997년	2,313건
ST	10,541건	18년(17.89년)	1976년~1990년	3,742건
계	76,187건			21,675건

〈표 3〉 각종 변수 및 데이터 산출방법

구분	변수	데이터 산출방법	
기본사항	등록번호	다운로드 데이터를 그대로 사용	
	등록일자	다운로드 데이터를 그대로 사용	
종속변수	피인용횟수*	다운로드 데이터를 토대로, 주제분야별 피인용반감기를 적용하여 산출	
독립 변수	형태적 요인	출원인 국적*	다운로드 데이터를 토대로 산출 [1=미국, 0=기타]
		출원인 유형*	다운로드 데이터를 토대로 산출 [1=기관, 0=개인]
		발명자 수*	다운로드 데이터를 토대로 산출
		출원국가 수*	다운로드 데이터를 토대로 산출(단, 동일 국가의 데이터가 2개 이상인 경우 1개 국가로 처리)
		페이지 수	수작업을 통해 해당 특허의 원문 페이지 수를 일일이 확인하여 산출
	기술적 요인	청구항 수	다운로드 데이터를 그대로 사용
		청구항 길이*	청구1항 길이의 byte 수를 계산하여 산출
		IPC 수*	다운로드 데이터를 토대로 산출
		참고문헌 수	'미국특허 인용정보'를 토대로 산출
		참고문헌 평균 인용시차*	분석대상 특허의 등록년도에서 각 참고문헌의 등록년도를 뺀 후, 이들 합을 참고문헌 수로 나누어 산출
		참고문헌 평균 피인용횟수*	분석대상 특허의 각 참고문헌에 대한 피인용수를 계산한 후, 이들 합을 참고문헌 수로 나누어 산출
		기술분야 피인용반감기	주제분야별 세부 기술분야(대표 IPC의 메인그룹단위)의 피인용반감기를 산출
	기술분야 특허집단 크기	주제분야별 세부 기술분야(대표 IPC의 메인그룹단위)의 특허등록건수를 산출	
	기술분야 특허증감율	주제분야별 세부 기술분야(대표 IPC의 메인그룹단위)의 구간별 특허증감율을 산출	
	개념적 요인	서지결합도*	분석대상 집단의 2개의 특허가 공통으로 인용한 특허의 건수를 산출
		동시인용도*	분석대상 집단의 2개의 특허를 동시에 인용한 특허의 건수를 산출
		문헌간유사도*	특허의 제목과 초록에 출현하는 용어를 대상으로 코사인 유사계수를 이용하여 산출

주) 별표(*)의 변수는 별도로 설계한 프로그램을 통해 변수 값을 산출한 것임.

로 서로 관계가 있는 2개 특허 간의 결합쌍 단위로 산출된 서지결합도, 동시인용도 및 문헌간유사도를 토대로, 분석대상 특허와 쌍결합된 대응특허의 피인용횟수를 독립변수의 변수값으로 사용하였다.

2.3 연구가설

이 연구의 목적은 최적의 특허인용 예측모형을 개발하기 위한 기초적 자료를 제공하고자 특

허의 피인용횟수에 어떠한 변수들이 얼마만큼의 상관관계를 갖고 있는지를 종합적으로 분석하는 것이다. 여기에서는 특허의 피인용횟수를 종속변수로 하고, 종속변수에 영향을 미칠 수 있는 제반 독립변수를 형태적 요인, 기술적 요인, 개념적 요인 등 3가지 측면으로 나누어 분석하였다.

형태적 요인은 출원인 국적, 출원인 유형, 발명자 수, 출원국가 수, 페이지 수 등 5가지 변수를 선정하였고, 기술적 요인은 청구항 수, 청구항 길이, IPC 수, 참고문헌 수, 참고문헌 평균

인용시차, 참고문헌 평균 피인용횟수, 기술분야 피인용반감기, 기술분야 특허집단 크기, 기술분야 특허증감율 등 9가지 변수를 선정하였으며, 개념적 요인은 서지결합도, 동시인용도, 문헌간 유사도 등 3가지 변수를 선정하였다.

이들 각각의 독립변수가 종속변수인 특허의 피인용횟수에 얼마간의 상관관계가 있는지를 검증하기 위해서 <표 4>와 같이 17가지의 연구 가설을 설정하였다. 참고로, 독립변수 중 더미 변수인 '출원인 국적'과 '출원인 유형'을 제외한 나머지 모든 독립변수와 종속변수인 피인용횟수는 사회과학분야에서 가장 보편적으로 사용되는 순위척도인 리커트 7점 척도를 사용하여 실제값을 7개 구간으로 나누어 분석하였다. 각 구간의 설정은 각 변수별로 해당 변수값의 분포비율과 변수의 특성을 충분히 고려하여 임의로 조정하였다.

3. 가설검증 및 분석

특허를 인용함에 있어서 어떠한 독립변수들이 종속변수에 얼마만큼의 상관관계가 있는지를 종합적으로 살펴보기 위해서 5개 주제분야에 걸쳐 총 21,675건의 특허를 분석대상으로 분석하였다. 이 연구에서는 연구가설을 검증하기 위해 SPSS for Window 12.0을 사용하여 상관관계분석을 실시하였다.

3.1 형태적 요인

특허의 피인용횟수와 형태적 요인의 5가지 변수들 간의 상관관계를 분석한 결과는 <표 5>와 같다.

첫째, '출원인 국적'은 ET분야를 제외한 4개의 주제분야와 5개 주제분야를 통합한 전체분야

<표 4> 연구가설

가 설	내 용
형태적 요인	H1 출원인 국적은 특허의 피인용횟수와 관련이 있을 것이다.
	H2 출원인 유형은 특허의 피인용횟수와 관련이 있을 것이다.
	H3 발명자 수는 특허의 피인용횟수와 관련이 있을 것이다.
	H4 출원국가 수는 특허의 피인용횟수와 관련이 있을 것이다.
	H5 페이지 수는 특허의 피인용횟수와 관련이 있을 것이다.
기술적 요인	H6 청구항 수는 특허의 피인용횟수와 관련이 있을 것이다.
	H7 청구항 길이는 특허의 피인용횟수와 관련이 있을 것이다.
	H8 IPC 수는 특허의 피인용횟수와 관련이 있을 것이다.
	H9 참고문헌 수는 특허의 피인용횟수와 관련이 있을 것이다.
	H10 참고문헌 평균 인용시차는 특허의 피인용횟수와 관련이 있을 것이다.
	H11 참고문헌 평균 피인용횟수는 특허의 피인용횟수와 관련이 있을 것이다.
	H12 기술분야 피인용반감기는 특허의 피인용횟수와 관련이 있을 것이다.
	H13 기술분야 특허집단 크기는 특허의 피인용횟수와 관련이 있을 것이다.
	H14 기술분야 특허증감율은 특허의 피인용횟수와 관련이 있을 것이다.
개념적 요인	H15 2개 특허간의 서지결합도는 특허의 피인용횟수와 관련이 있을 것이다.
	H16 2개 특허간의 동시인용도는 특허의 피인용횟수와 관련이 있을 것이다.
	H17 2개 특허간의 문헌간유사도는 특허의 피인용횟수와 관련이 있을 것이다.

〈표 5〉 피인용횟수와 형태적 요인 변수들 간의 상관관계분석 결과

		출원인 국적	출원인 유형	발명자 수	출원국가 수	페이지 수
BT	r(p)	-.201(.000)	-.013(.350)	-.048(.000)	.140(.000)	.257(.000)
	r ²	.040	.000	.002	.020	.066
ET	r(p)	-.001(.987)	-.042(.001)	.132(.000)	.101(.000)	.234(.000)
	r ²	.000	.002	.017	.010	.055
IT	r(p)	-.135(.000)	-.009(.568)	.171(.000)	.033(.028)	.266(.000)
	r ²	.018	.000	.029	.001	.071
NT	r(p)	-.109(.000)	-.136(.085)	.139(.000)	.172(.000)	.287(.000)
	r ²	.012	.001	.019	.030	.082
ST	r(p)	-.063(.000)	-.105(.000)	.101(.000)	.067(.000)	.162(.000)
	r ²	.004	.011	.010	.005	.026
전체	r(p)	-.081(.000)	-.060(.000)	.072(.000)	.094(.000)	.278(.000)
	r ²	.007	.004	.005	.009	.077

주) 표 안의 r(p)는 피어슨 상관계수(유의확률)이며, r²은 설명 변량임.

에서 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 있는 것으로 나타났다(r=-.201~-.063, p<.05). 따라서 가설-1은 부분적으로 채택되었다. 유의미한 주제분야에서의 '출원인 국적'이 특허의 피인용횟수를 설명할 수 있는 변량(r²)은 .004~.040으로 약 0.4%~4% 정도의 설명력을 갖는 것으로 나타났다.

둘째, '출원인 유형'은 ET, ST분야와 전체분야에서만 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 있는 것으로 나타났다(r=-.042~-.105, p<.05). 따라서 가설-2도 부분 채택되었다. 유의미한 주제분야에서의 '출원인 유형'이 특허의 피인용횟수를 설명할 수 있는 변량은 .002~.011로 약 0.2%~1.1% 정도의 설명력을 갖는 것으로 나타났다.

셋째, '발명자 수'는 모든 주제분야에서 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 있는 것으로 나타났다(r=-.048~.171, p<.05). 따라서 가설-3은 채택되었다. '발명자 수'가 특허의 피인용횟수를 설명할 수 있는 변량은 .002~

.029로 약 0.2%~2.9% 정도의 설명력을 갖는 것으로 나타났다.

넷째, '출원국가 수' 역시 모든 주제분야에서 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 있는 것으로 나타났다(r=.033~.172, p<.05). 따라서 가설-4는 채택되었다. '출원국가 수'가 특허의 피인용횟수를 설명할 수 있는 변량은 .005~.030으로 약 0.5%~3.0% 정도의 설명력을 갖는 것으로 나타났다.

다섯째, '페이지 수'도 모든 주제분야에서 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 있는 것으로 나타났다(r=.162~.287, p<.05). 따라서 가설-5는 채택되었다. '페이지 수'가 특허의 피인용횟수를 설명할 수 있는 변량은 .026~.082로 약 2.6%~8.2% 정도의 설명력을 갖는 것으로 나타났다.

3.2 기술적 요인

특허의 피인용횟수와 기술적 요인의 9가지

변수들 간의 상관관계를 분석한 결과는 <표 6>과 같다.

첫째, '청구항 수'는 모든 주제분야에서 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 있는 것으로 나타났다($r=.154\sim.247, p<.05$). 따라서 가설-6은 채택되었다. '청구항 수'가 특허의 피인용횟수를 설명할 수 있는 변량(r^2)은 .024~.067로 약 2.4%~6.7% 정도의 설명력을 갖는 것으로 나타났다.

둘째, '청구항 길이'는 BT, ET분야와 전체분야에서만 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 있는 것으로 나타났다($r=.067\sim.147, p<.05$). 따라서 가설-7은 부분적으로만 채택되었다. 유의미한 주제분야에서의 '청구항 길이'가 특허의 피인용횟수를 설명할 수 있는 변량은 .004~.021로 약 0.4%~2.1% 정도의 설명력을 갖는 것으로 나타났다.

셋째, 'IPC 수'는 BT, ET, IT, ST분야에서만 특허의 피인용횟수와 통계적으로 유의미한 상

관관계가 있는 것으로 나타났다($r=.039\sim.083, p<.05$). 따라서 가설-8은 부분적으로만 채택되었다. 유의미한 주제분야에서의 'IPC 수'가 특허의 피인용횟수를 설명할 수 있는 변량은 .001~.007로 약 0.1%~0.7% 정도의 설명력을 갖는 것으로 나타났다.

넷째, '참고문헌 수'는 모든 주제분야에서 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 있는 것으로 나타났다($r=.103\sim.219, p<.05$). 따라서 가설-9는 채택되었다. '참고문헌 수'가 특허의 피인용횟수를 설명할 수 있는 변량은 .011~.048로 약 1.1%~4.8% 정도의 설명력을 갖는 것으로 나타났다.

다섯째, '참고문헌 평균 인용시차' 역시 모든 주제분야에서 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 있는 것으로 나타났다($r=-.167\sim.126, p<.05$). 따라서 가설-10은 채택되었으며, '참고문헌 평균 인용시차'가 특허의 피인용횟수를 설명할 수 있는 변량은 .009~

<표 6> 피인용횟수와 기술적 요인 변수들 간의 상관관계분석 결과

		청구항 수	청구항 길이	IPC 수	참고문헌 수	참고문헌 평균 인용시차	참고문헌 평균 피인용수	기술분야 피인용반 감기	기술분야 특허집단 크기	기술분야 특허 증감율
BT	r(p)	.239(.000)	.094(.000)	.078(.000)	.126(.000)	-.097(.000)	.409(.000)	.566(.144)	-.343(.406)	.751(.000)
	r ²	.057	.009	.006	.016	.009	.167	.320	.117	.563
ET	r(p)	.180(.000)	.067(.000)	.083(.000)	.103(.000)	.126(.000)	.094(.000)	-.033(.922)	-.560(.073)	.856(.000)
	r ²	.032	.004	.007	.011	.016	.009	.001	.313	.733
IT	r(p)	.247(.000)	-.022(.151)	.059(.000)	.177(.000)	-.154(.000)	.468(.000)	.414(.181)	-.045(.890)	.626(.000)
	r ²	.061	.000	.003	.031	.024	.219	.171	.002	.392
NT	r(p)	.260(.000)	.019(.353)	.001(.958)	.219(.000)	-.122(.000)	.496(.000)	.409(.000)	.002(.984)	.277(.000)
	r ²	.067	.000	.000	.048	.015	.246	.168	.000	.077
ST	r(p)	.154(.000)	.019(.250)	.039(.018)	.147(.000)	-.137(.000)	.356(.000)	-.066(.680)	.095(.549)	.552(.000)
	r ²	.024	.000	.001	.022	.019	.127	.004	.009	.305
전체	r(p)	.239(.000)	.147(.000)	-.005(.490)	.118(.000)	-.167(.000)	.228(.000)	.117(.120)	-.008(.912)	.306(.000)
	r ²	.057	.021	.000	.014	.028	.052	.014	.000	.094

주) 표 안의 r(p)는 피어슨 상관계수(유의확률)이며, r²은 설명 변량임.

.028로 약 0.9%~2.8% 정도의 설명력을 갖는 것으로 나타났다.

여섯째, '참고문헌 평균 피인용횟수'도 모든 주제분야에서 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 있는 것으로 나타났다($r=.094\sim.496$, $p<.05$). 따라서 가설-10은 채택되었다. '참고문헌 평균 피인용횟수'가 특허의 피인용횟수를 설명할 수 있는 변량은 .009~.246로 약 0.9%~24.6% 정도의 설명력을 갖는 것으로 나타났다.

일곱째, '기술분야 피인용반감기'는 NT분야에서만 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 있는 것으로 나타났다($r=.409$, $p<.05$). 따라서 가설-12는 부분 채택되었다. NT분야에서 '기술분야 피인용반감기'가 특허의 피인용횟수를 설명할 수 있는 변량은 .409로 약 40.9% 정도의 설명력을 갖는 것으로 나타났다.

여덟째, '기술분야 특허집단 크기'는 모든 주제분야에 있어서 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 없는 것으로 나타났다($p<.05$). 따라서 가설-13은 기각되었다.

아홉째, '기술분야 특허증감율'은 모든 주제분야에서 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 있는 것으로 나타났다($r=.277\sim.856$, $p<.05$). 따라서 가설-14는 채택되었다. '기술분야 특허증감율'이 특허의 피인용횟수를 설명할 수 있는 변량은 .077~.733로 약 7.7%~73.3% 정도의 설명력을 갖는 것으로 나타났다.

3.3 개념적 요인

특허의 피인용횟수와 개념적 요인의 3가지

변수들 간의 상관관계를 분석한 결과는 <표 7>과 같다.

첫째, 서지결합도는 BT, ET, IT분야와 전체분야에서 모든 구간값에 있어서 통계적으로 유의미한 상관관계가 있는 것으로 나타난 반면, NT분야에서는 1~2구간에서만, ST분야에서는 1~4구간에서만 통계적으로 유의미한 것으로 나타났다($p<.05$). 따라서 가설-15는 BT, ET, IT분야와 전체분야에서만 채택되었고, NT, ST분야에서는 부분적으로만 채택되었으며, 전반적인 측면에서 보면 서지적으로 결합된 2개 특허 간에 서지결합도가 높아질수록 그들 간의 피인용횟수에 대한 상관관계수가 증가하는 것으로 나타났다. 서지결합도가 특허의 피인용횟수를 설명할 수 있는 변량(r^2)은 구간값에 따라 BT분야 .026~.834, ET분야 .052~.361, IT분야 .018~.464 그리고 전체분야는 .079~.511로 나타나 각각 약 2.6%~83.4%, 5.2%~36.1%, 1.8%~46.4%, 7.9%~51.1% 정도의 설명력을 갖는 것으로 나타났으며, NT분야와 ST분야는 유의미한 구간에서의 변량이 구간값에 따라 .074~.123과 .021~.091로 나타나 각각 7.4%~12.3%와 2.1%~9.1% 정도의 설명력을 갖는 것으로 나타났다.

둘째, 동시인용도는 5개 주제분야에서는 물론 전체분야에서도 모든 구간값에 있어서 통계적으로 유의미한 상관관계가 있는 것으로 나타났다($p<.05$). 따라서 가설-16은 모든 주제분야에서 채택되었는데, 전반적으로 모든 주제분야에 있어서 2개 특허 간의 동시인용도의 구간값에 따른 상관관계수의 증감율이 일정하지 않은 것으로 나타났다. 동시인용도가 특허의 피인용횟수를 설명할 수 있는 변량은 구간값에 따라

<표 7> 피인용횟수와 개념적 요인 변수들 간의 상관관계분석 결과

		구간값	BT	ET	IT	NT	ST	전체
서지 결합도	1	r(p)	.162(.000)	.227(.000)	.134(.000)	.273(.000)	.181(.000)	.280(.000)
		r ²	.026	.052	.018	.074	.033	.079
	2	r(p)	.249(.000)	.328(.000)	.289(.000)	.351(.000)	.234(.000)	.375(.000)
		r ²	.062	.107	.084	.123	.055	.140
	3	r(p)	.464(.000)	.444(.000)	.309(.000)	.251(.059)	.144(.029)	.448(.000)
		r ²	.216	.197	.095	.063	.021	.200
	4	r(p)	.475(.000)	.372(.000)	.429(.000)	.783(.117)	.302(.003)	.426(.000)
		r ²	.225	.139	.184	.613	.091	.182
	5	r(p)	.754(.000)	.481(.000)	.681(.000)	-	.343(.059)	.568(.000)
		r ²	.569	.232	.464	-	.118	.323
	6	r(p)	.848(.001)	.500(.000)	.674(.003)	-	.396(.061)	.600(.000)
		r ²	.719	.250	.454	-	.157	.360
	7이상	r(p)	.913(.000)	.601(.000)	.630(.000)	-	-.089(.795)	.715(.000)
		r ²	.834	.361	.397	-	.008	.511
동시 인용도	1	r(p)	.154(.000)	.141(.000)	.156(.000)	.167(.000)	.105(.000)	.318(.000)
		r ²	.024	.020	.024	.028	.011	.101
	2~3	r(p)	.173(.000)	.167(.000)	.119(.000)	.211(.000)	.196(.000)	.330(.000)
		r ²	.030	.028	.014	.045	.038	.109
	4~5	r(p)	.134(.000)	.243(.000)	.156(.000)	.130(.000)	.349(.000)	.380(.000)
		r ²	.018	.059	.024	.017	.122	.145
	6~10	r(p)	.100(.000)	.144(.000)	.183(.000)	.150(.000)	.394(.000)	.311(.000)
		r ²	.010	.021	.033	.023	.155	.097
	11~20	r(p)	.162(.000)	.337(.000)	.339(.000)	.148(.000)	.155(.000)	.446(.000)
		r ²	.026	.114	.115	.022	.024	.199
	21~30	r(p)	.287(.000)	.472(.000)	.238(.000)	.334(.000)	.714(.020)	.343(.000)
		r ²	.083	.222	.057	.112	.510	.118
	31이상	r(p)	.040(.000)	.386(.000)	.288(.000)	.539(.000)	1.000(.000)	.412(.000)
		r ²	.002	.149	.083	.290	1.00	.170
문헌 간유사도	0.30~0.39	r(p)	.251(.000)	.216(.000)	.217(.000)	.277(.000)	.117(.000)	.328(.000)
		r ²	.063	.047	.047	.077	.014	.108
	0.40~0.49	r(p)	.365(.000)	.275(.000)	.249(.000)	.448(.000)	.191(.000)	.404(.000)
		r ²	.133	.076	.062	.201	.037	.163
	0.50~0.59	r(p)	.387(.000)	.400(.000)	.332(.000)	.442(.000)	.296(.000)	.474(.000)
		r ²	.150	.160	.110	.195	.088	.224
	0.60~0.69	r(p)	.475(.000)	.446(.000)	.321(.000)	.635(.000)	.236(.000)	.515(.000)
		r ²	.226	.199	.103	.403	.056	.265
	0.70~0.79	r(p)	.579(.000)	.473(.000)	.651(.000)	.790(.000)	.436(.000)	.645(.000)
		r ²	.335	.224	.424	.623	.190	.416
	0.80~0.89	r(p)	.741(.000)	.572(.000)	.766(.000)	.841(.000)	.446(.006)	.740(.000)
		r ²	.549	.327	.587	.707	.198	.547
	0.90이상	r(p)	.652(.000)	.592(.000)	.834(.000)	.689(.000)	.480(.000)	.680(.000)
		r ²	.425	.351	.695	.475	.230	.463

주) 표 안의 r(p)는 피어슨 상관계수(유의확률)이며, r²은 설명 변량임.

BT분야 .002~.083, ET분야 .020~.222, IT분야 .014~.115, NT분야 .017~.290, ST분야 .011~1.000 그리고 전체분야는 .097~.199로 나타나 각각 0.2%~8.3%, 2.0%~22.2%, 1.4%~11.5%, 1.7%~29.0%, 1.1%~100%, 9.7%~19.9% 정도의 설명력을 갖는 것으로 나타났다.

셋째, 문헌간유사도 역시 5개 주제분야에서는 물론 전체분야에서도 모든 구간값에 있어서 통계적으로 유의미한 상관관계가 있는 것으로 나타났다($p < .05$). 따라서 가설-17은 모든 주제분야에서 채택되었으며, 전반적으로 2개 특허간의 문헌간유사도가 높아질수록 그들 간의 피인용횟수에 대한 상관계수가 점점 증가하는 것으로 나타났다. 문헌간유사도가 특허의 피인용횟수를 설명할 수 있는 변량은 구간값에 따라 BT분야 .063~.549, ET분야 .047~.592, IT분야 .047~.695, NT분야 .077~.841, ST분야 .014~.230 그리고 전체분야는 .108~.547로 나타나 각각 6.3%~54.9%, 4.7%~59.2%, 4.7%~69.5%, 7.7%~84.1%, 1.4%~23.0%, 10.8%~54.7% 정도의 설명력을 갖는 것으로 나타났다.

이상에서의 상관관계분석 결과를 종합 정리하면, 분석대상인 17개의 변수 가운데 '기술분야 특허집단 크기'를 제외한 16개의 변수가 특허의 피인용횟수와 통계적으로 유의미한 상관관계가 있는 것으로 밝혀졌다. 그러나 5개 주제분야를 통합한 전체분야를 토대로 특허의 피인용횟수를 설명할 수 있는 변량을 살펴보면, 피인용횟수와 일정 수준 이상의 상관관계, 즉 변량이 .05 이상으로 5% 이상의 설명력이 있는 변수는 페이지 수, 청구항 수, 참고문헌 평균 피인용횟수, 기술분야 특허증감율, 서지결합도, 동시인용도, 문헌간유사도 등 7개에 불과한 것

으로 나타났다.

4. 주제분야간 동질성 분석

특허의 인용에 일정 수준 이상의 상관관계가 있는 제반 변수들이 각 집단, 즉 주제분야 별로 차이가 있는지를 알아보기 위해서 SPSS for window 12.0을 사용하여 일원분산분석(one-way ANOVA)을 실시하였다. 분산분석 검증결과 유의한 결과를 보인 변수, 즉 집단 간의 차이가 있는 변수에 대해서는 어느 집단 간에 차이가 있는지를 확인하기 위해 가장 일반적으로 사용되는 방법 중의 하나인 Scheffe 방법을 사용하여 사후분석을 실시하였다.

여기에서의 분석대상 변수는 앞의 상관관계 분석에서 특허의 피인용횟수와 일정 수준 이상의 상관관계, 즉 변량(r^2)이 .05 이상으로 5% 이상의 설명력이 있는 7개의 변수만을 그 대상으로 선정하였다. 이들 변수에 대한 주제분야별 분산분석 결과 및 Scheffe 사후검증에 따른 주제분야간 평균차를 정리하면 <표 8>, <표 9>와 같다.

참고로, 개념적 요인의 3가지 변수의 경우 앞 장의 상관관계분석에서와는 달리 특허쌍 단위가 아닌 개별 특허 단위로 해당 변수값을 산출해야 하기 때문에 서지결합도, 동시인용도 및 문헌간유사도를 가중치로 사용하여 분석대상 특허와 쌍결합된 대응특허의 피인용횟수에 그 가중치를 곱한 값의 평균값을 해당 변수의 값으로 사용하였다. 일례로, 서지결합도의 변수 값 산출과정을 보면, 분석대상 특허와 쌍결합을 이루는 각 대응특허의 피인용횟수에 해당

〈표 8〉 주제분야별 주요 변수에 대한 분산분석 결과

		(n)	$\bar{x} \pm S.D.$	$F(p)$
피인용횟수	BT	5,337	8.36 ± 13.01	499,460(.000)
	ET	5,929	7.74 ± 9.93	
	IT	4,354	19.45 ± 26.66	
	NT	2,313	12.41 ± 16.12	
	ST	3,742	6.29 ± 6.60	
페이지수	BT	5,337	16.50 ± 15.50	642,236(.000)
	ET	5,929	7.90 ± 4.64	
	IT	4,354	17.20 ± 18.04	
	NT	2,313	10.60 ± 7.62	
	ST	3,742	8.92 ± 5.20	
청구항 수	BT	5,337	11.49 ± 11.50	76,664(.000)
	ET	5,929	12.41 ± 9.46	
	IT	4,354	13.94 ± 12.00	
	NT	2,313	14.43 ± 12.00	
	ST	3,742	10.68 ± 8.90	
참고문헌 평균 피인용횟수	BT	3,838	45.85 ± 102.63	387,859(.000)
	ET	4,365	9.31 ± 14.90	
	IT	4,328	35.62 ± 27.97	
	NT	2,247	29.34 ± 29.53	
	ST	3,726	11.44 ± 8.78	
기술분야 특허증감율	BT	30	1.59 ± 1.82	12,110(.000)
	ET	41	0.29 ± 1.12	
	IT	43	1.25 ± 1.96	
	NT	320	2.05 ± 3.54	
	ST	167	0.37 ± 1.12	
서지결합도	BT	1,439	13.03 ± 11.86	537,318(.000)
	ET	3,668	8.71 ± 6.59	
	IT	2,892	21.04 ± 17.50	
	NT	252	21.53 ± 16.51	
	ST	1,217	7.77 ± 4.74	
동시결합도	BT	3,419	18.54 ± 16.39	1,726,395(.000)
	ET	5,186	16.60 ± 10.94	
	IT	3,866	44.74 ± 33.04	
	NT	1,170	23.76 ± 18.63	
	ST	3,018	10.15 ± 6.30	
문헌간유사도	BT	2,924	11.30 ± 14.13	547,641(.000)
	ET	4,221	7.61 ± 6.89	
	IT	3,244	18.87 ± 16.74	
	NT	1,001	14.74 ± 16.74	
	ST	2,523	6.31 ± 4.27	

〈표 9〉 Scheffe 사후검증에 따른 주제분야간 평균차(I-J) 비교

(I) 분류	BT				ET			IT		NT
(J) 분류	ET	IT	NT	ST	IT	NT	ST	NT	ST	ST
피인용횟수	0.617	-11.090*	-4.052*	2.065*	-11.707*	-4.669*	1.448*	7.038*	13.156*	6.118*
페이지 수	8.603*	-0.696	5.905*	7.587*	-9.299*	-2.698*	-1.016*	6.600*	8.283*	1.682*
청구항 수	-0.919*	-2.448*	-2.934*	0.810*	-1.529*	-2.015*	1.729*	-0.485	3.258*	3.744*
참고문헌 평균 피인용횟수	36.531*	10.223*	16.504*	34.408*	-26.308*	-20.027*	-2.123	6.281*	24.185*	17.905*
기술분야 특허증감율	1.299	0.338	-0.464	1.222	-0.960	-1.762*	-0.077	-0.802	0.884	1.686*
서지결합도	4.319*	-8.014*	-8.502*	5.261*	-12.333*	-12.821*	0.942	-0.488	13.275*	13.763*
동시인용도	1.942*	-26.194*	-5.216*	8.393*	-28.136*	-7.158*	6.451*	20.978*	34.587*	13.609*
문헌간유사도	3.687*	-7.573*	-3.440*	4.987*	-11.260*	-7.128*	1.300*	4.132*	12.560*	8.428*

* .05 수준에서 유의한 평균차가 있다.

특허쌍의 서지결합도를 곱한 후, 이들의 합을 서지결합도의 합으로 나누어 산출하였다. 동시 인용도와 문헌간유사도의 변수값 산출과정도 마찬가지이다.

먼저, 상관관계분석에서의 유일한 종속변수인 특허 피인용횟수에 대한 분산분석 결과를 정리하면 다음과 같다.

각 집단, 즉 주제분야 간의 특허 피인용횟수에 대한 F-통계량은 499.460으로 나타났다. F-분포에서의 유의수준 .05에서 F의 임계치가 2.37인데 임계치보다 F-통계량이 크므로 각 집단 간의 평균이 동일하지 않다는 대립가설이 채택되었다.

이는 Scheffe 사후분석을 통해서도 각 집단 간의 유의한 평균차를 검증할 수 있다. 즉 유의수준 .05에서 BT-ET분야 간의 경우를 제외한 모든 주제분야 간에는 특허 피인용횟수의 평균에 있어서 통계적으로 유의미한 평균차가 있는 것으로 밝혀졌다. 이로써 보면, 특허의 피인용횟수의 평균은 각 주제분야에 따라 차이가 있다는 결론을 내릴 수 있다.

다음, 상관관계분석 결과 특허의 피인용횟수와 일정 수준 이상의 상관관계가 있는 7개의 변수에 대한 분산분석 결과를 정리하면 다음과 같다.

7개 변수에 대한 각 집단, 즉 주제분야 간의 F-통계량을 보면, 페이지 수가 642.236, 청구항 수 76.664, 참고문헌 평균 피인용횟수 387.859, 기술분야 특허증감율 12.110, 서지결합도 537.318, 동시인용도 1,726.395 그리고 문헌간유사도가 547.641로 나타났다. 7개의 변수 모두 F-분포에서의 유의수준 .05에서 F의 임계치가 2.37인데, 이들 변수 모두 임계치보다 F-통계량이 크므로 각 집단 간의 평균이 동일하지 않다는 대립가설이 채택되었다.

Scheffe 사후분석을 통해 각 집단 간의 평균차를 검증한 결과에 의하면, 유의수준 .05에서 페이지 수는 BT-IT분야 간, 청구항 수는 IT-NT분야 간, 참고문헌 평균 피인용횟수 ET-ST분야 간을 제외한 모든 주제분야 간에 있어서 각 변수의 평균에 있어서 통계적으로 유의미한 평균차가 있는 것으로 나타났으며, 동시인용도와 문헌간유사도는 모든 주제분야 간에 있어서 통계적으로 유의미한 평균차가 있는 것으로 나타났다. 반면 기술분야 특허증감율은 오직 ET-NT분야와 NT-ST분야에서만 통계적으로 유의미한 평균차가 있는 것으로 나타났다. 이로써 보면, 특허의 피인용횟수와 일정 수준 이상의 상관관

계가 있는 7개 변수에 대한 평균은 각각 주제분야에 따라 차이가 있다는 결론을 내릴 수 있다.

이상에서의 분산분석 검증결과를 종합 정리하면, 특허의 피인용횟수는 물론 이에 일정 수준 이상의 상관관계가 있는 7개 변수들 모두 각 집단, 즉 주제분야 간에 전반적으로 평균의 차이가 있음을 알 수 있다. 따라서 향후 이를 토대로 특허인용 예측모형을 설계할 경우 각 주제분야별로 나누어 예측모형을 추정하는 것이 바람직할 것으로 판단된다. 반면, 일부 변수의 경우 특정 주제분야 간에 있어서 변수값의 평균차가 없는 것으로 나타났으므로 이들 주제분야는 하나로 통합하여 모형을 추정하는 것도 적극 검토할 필요가 있을 것으로 보인다.

5. 결 론

이 연구에서는 특허의 인용에 영향을 미칠 수 있는 제반 요인들을 종합적으로 살펴보고자 5개 주제분야를 대상으로 형태적 요인, 기술적 요인, 개념적 요인 등 3가지 측면으로 나누어 분석하였다. 그 결과 어떠한 변수들이 특허의 피인용횟수와 상관관계가 있는지를 찾아내고 그러한 변수값들이 주제분야별로 차이가 있는지를 검토하였다.

특허의 피인용횟수에 어떠한 변수들이 얼마만큼의 관련이 있는지를 확인하고자 실시한 상관관계분석 결과는 다음과 같다.

첫째, 형태적 요인의 경우 페이지 수만 특허의 피인용횟수와 낮은 상관관계가 있는 것으로 나타났으며, 출원인 국적, 출원인 유형, 발명자 수 및 출원국가 수는 상관관계가 거의 없는 것

으로 나타났다.

둘째, 기술적 요인의 경우 청구항 수, 참고문헌 평균 피인용횟수 및 기술분야 특허증감율은 특허의 피인용횟수와 주제분야에 따라서 낮거나 보통 수준의 상관관계가 있는 것으로 나타난 반면, 청구항 길이, IPC 수, 참고문헌 수, 참고문헌 평균 인용시차 및 기술분야 피인용반감기는 상관관계가 거의 없는 것으로 나타났으며, 기술분야 특허집단 크기는 전혀 상관관계가 없는 것으로 나타났다.

셋째, 개념적 요인의 경우 서지결합도와 문헌간유사도는 변수값이 증가할수록 특허의 피인용횟수와 상관관계가 높아지는 반면, 동시인용도는 주제분야는 물론이고 변수값에 따라서도 특허의 피인용횟수와 상관관계가 일정하지 않은 것으로 나타났다.

위의 상관관계분석 결과를 토대로 특허의 피인용횟수와 일정 수준 이상의 상관관계가 있는 7개의 변수들을 대상으로 이들 변수값의 평균이 주제분야 간에 차이가 있는지를 확인하고자 분산분석을 실시하였다.

그 결과 페이지 수, 청구항 수, 참고문헌 평균 피인용횟수, 기술분야 특허증감율 및 서지결합도는 일부 주제분야 간에 있어서만 변수값의 평균차가 없을 뿐 나머지 주제분야간에는 모두 평균차가 있는 것으로 나타났다. 반면 동시인용도와 문헌간유사도의 경우에는 5개 주제분야 간에 있어서 모두 변수값의 평균차가 있는 것으로 나타났다.

이 연구는 특허의 인용에 영향을 미칠만한 형태적·기술적·개념적 측면의 다양한 변수들을 종합적으로 분석하였다는 데 그 의미가 있다. 향후에는 이 연구에서 도출된 특허의 피

인용횟수에 일정 수준 이상 상관관계가 있는 소위 예측변수들을 토대로 실제로 예측모형을

설계하여 예측값과 실제값 간의 차이를 비교해 보는 것도 필요할 것으로 보인다.

참 고 문 헌

- 노경란. 2006. 『특허분석을 통한 과학기술자의 과학논문 인용행태에 관한 연구』. 박사학위논문, 연세대학교 대학원, 문헌정보학과.
- Dutta, S., and Weiss, A. M. 1997. "The relationship between a firm's level of technological innovativeness and its pattern of partnership agreements." *Management Science*, 43(3): 343-356.
- Hall, B. H., A. B. Jaffe, and M. Trajtenberg. 2001. "The NBER patent citation data file: lessons, insights and methodological tools," *NBER Working Paper No. 8498*, Cambridge: National Bureau of Economic Research.
- Iversen, E. J. 2000. "An excursion into the patent-bibliometrics of Norwegian patenting." *Scientometrics*, 49(1): 63-80.
- Lanjouw, J. O., and M. Schankerman. 2004. "Patent quality and research productivity: measuring innovation with multiple indicators." *The Economic Journal*, 114(495): 441-465.
- Lee, Yong-Gil, Jeong-Dong Lee, and Yong-Il Song. 2006. "An Analysis of citation counts of ETRI-invented US patents." *ETRI Journal*, 28(4): 541-544.
- Lin, Bou-Wen, Chung-Jen Chen, and Hsueh-Liang Wu. 2007. "Predicting citations to biotechnology patents based on the information from the patent documents." *International Journal of Technology Management*, 40(1-3): 87-100.
- Moorman, C., and R. J. Slotegraaf. 1999. "The contingency value of complementary capabilities in product development." *Journal of Marketing Research*, 36(2): 239-257.
- Narin, F., M. P. Carpenter, and P. Woolf. 1984. "Technological performance assessments based on patents and patent citations." *IEEE Transactions on Engineering Management*, 36(2): 172-183.
- Narin, F., K. S. Hamilton, and D. Olivastro. 1997. "The increasing linkage between US technology and public science." *Research Policy*, 26(3): 317-330.
- Trajtenberg, M. 1990. "A penny for your quotes: Patent citations and the value of innovations." *The Rand Journal of Economics*, 21(1): 172-187.