(19)대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) 。Int. Cl.[8]
G06F 17/00 (2006.01)

(45) 공고일자 2006년02월02일
(11) 등록번호 10-0548983
(24) 등록일자 2006년01월26일

(21) 출원번호 10-2000-0065038
(22) 출원일자 2000년11월02일

(65) 공개번호 10-2001-0095343
(43) 공개일자 2001년11월07일

(73) 특허권자 (주)마크텍
서울 중구 쌍림동 151-11 쌍림빌딩 10층

주식회사 마크애니
서울 중구 쌍림동 151-11

(72) 발명자 최종욱
서울특별시강북구우이동1성원아파트2-1301

최기철
서울특별시종로구홍지동68-1

(74) 대리인 특허법인코리아나

심사관 : 박성우

(54) 디지털 증명서의 발급 및 인증을 위한 텍스트의 삽입 방법및 장치

요약

본 발명은 텍스트 임베딩(Text Embedding) 방법을 이용한 각종 디지털 증명서의 인증을 위한 방법 및 장치에 관한 것이다. 본 발명에 따르면 증명서가 위조 또는 변조되었을 경우, 그 변동 상황을 알아내고, 위조 및 변조된 부분을 검출하며, 변동 상황 검출과 함께 원본 증명서의 내용을 복원하는 방법과 기술 및 장치가 제공된다.

본 발명은 인증을 위한 일련의 정보를 담은 텍스트(이하 "정보 텍스트")가 삽입된 증명서를 발급하기 위한 장치 및 소프트웨어, 텍스트가 삽입된 증명서를 인식하기 위한 장치 및 소프트웨어가 포함된다. 본 발명의 실행은 증명서의 특성에 의하여 두 가지로 나뉘어 진다. 증명서 내용에 이미지 형태가 포함되었을 경우와 증명서에 이미지 형태가 포함되어 있지 않을 경우이다. 증명서에 이미지가 포함되는 경우, 증명서에 포함되어있는 이미지에 텍스트 임베딩 기법을 적용하여 발급자가 정하는 일련의 정보 텍스트(증명서 고유 번호나 발급기관, 담당자 이름 등)를 삽입하며, 증명서에 이미지가 포함되어있지 않을 경우는 증명서의 바탕색과 같은 색으로 되어 육안으로 식별이 불가능한 이미지(이하 "투명 이미지" : 보통 흰색 이미지)를 생성하고 삽입하고자 하는 정보 텍스트를 상기 투명 이미지에 삽입한 후 증명서에 다시 삽입한다. 삽입하는 정보 텍스트는 영문 텍스트를 기본으로 하여, 한국어, 중국어, 일본어를 포함한 2바이트 완성형 코드인 국제 표준인 ISO 10646 코드, Unicode를 포함한다. 디지털 증명서의 발급을 위한 텍스트 임베딩 방법은 증명서에 이미지가 포함될 경우는 세 단계 즉, 삽입할 정보 텍스트를 생성하는 단계, 생성한 정보 텍스트를 증명서에 포함되어 있는 이미지에 삽입하는 단계, 삽입한 정보 텍스트를 추출하여 증명서를 인증하는 단계로 이루어진다. 증명서에 이미지 형태의 파일이 포함되어 있지 않은 경우는 삽입할 정보 텍스트를 생성하는 단계, 투명 이미지를 생성하는 단계, 생성한 정보 텍스트를 투명 이미지에 삽입하여 증명서에 넣는 단계, 증명서의 인증을 위해 투명 이미지에서 정보 텍스트를 추출하는 단계로 이루어진다.

본 발명은 증명하려는 내용이 담긴 문서 형태를 기본으로 하여 작성되고 발급되는 각종 증명서의 위조 및 변조를 방지할 수 있으며, 증명서 제공자와 증명서 사용자의 권익을 보호할 수 있다.

대표도

도 1

색인어

전자문서, 인증, 나머지 연산자, 디지털 증명서,바 코드

명세서

도면의 간단한 설명

도 1은 본 발명의 텍스트 삽입 장치의 삽입 과정을 개념적으로 도시한 흐름도.

도 2는 본 발명의 텍스트 추출 과정을 개념적으로 도시한 흐름도

도 3는 각종 디지털 증명서의 종류를 예시한 도면 및 디지털 증명서의 특성 분석

도 4는 증명서의 특성에 근거하여 증명서를 분류하여 각 특성에 따른 삽입할 텍스트를 생성하는 단계로 가는 과정을 나타낸 흐름도

도 5는 영문 텍스트 문서를 기반으로 한 증명서에 있어 삽입할 텍스트의 비트열화 과정

도 6는 도 4의 분류에 기초하여 영문 외의 언어일 경우, 삽입할 텍스트를 비트열화 하는 과정을 도면화 한 것

도 7은 디지털 증명서에 포함되는 이미지 혹은 증명서에 이미지가 포함되어 있지 않을 경우 생성한 투명 이미지를 1차원 데이터로 바꾸어주는 과정을 나타낸 도면

도 8은 투명 이미지의 생성 과정을 나타낸 도면

도 9A는 텍스트 데이터의 삽입 과정을 나타낸 도면

도 9B 는 텍스트 데이터의 삽입 과정의 예를 나타낸 도면

도 10은 텍스트 데이터의 추출 과정을 나타낸 도면 도 11은 텍스트가 삽입되기 전의 단색 이미지와 텍스트가 삽입된 후의 단색 이미지를 히스토그램과 함께 비교해본 결과를 나타낸 도면

도 12는 텍스트가 삽입되기 전의 칼라 이미지와 텍스트가 삽입된 후의 칼라 이미지를 히스토그램과 함께 비교해본 결과를 나타낸 도면

도 13은 텍스트가 삽입되기 전의 인감 이미지와 텍스트가 삽입된 후의 인감 이미지를 히스토그램과 함께 비교해본 결과를 나타낸 도면

도 14는 증명서에 포함되는 이미지 데이터의 실례들을 나타낸 도면

도 15 는 바 코드 이미지를 포함한 증명서의 실례를 나타낸 도면

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 디지털 증명서의 특성에 근거하여 증명서의 위조 및 변조를 방지하고 증명서가 원본임을 인증할 수 있도록 하는 방법과 장치에 관한 것이다. 좀 더 상세하게는 디지털 증명서에 그 증명서를 인증하기 위한 일련의 정보가 담긴 텍스트를 보이지 않게 삽입한 후 증명서를 발급해 주고, 본 발명에서 제시되는 방법과 장치를 통해 삽입되어 있는 정보 텍스트의 추출 및 추출된 정보 텍스트의 내용 확인 등을 통해 디지털 증명서를 인증, 확인하는 것이다. 여기서 "삽입" 은 제 1 개체가 제 2 개체에 물리적으로 배치되는 의미 뿐 아니라 제 1 개체의 내용에 근거하여 제 2 개체의 내용을 특정 알고리즘을 사용하여 변경함으로써 나중에 사용된 알고리즘과 제 2 개체만을 가지고 제 1 개체의 내용을 추출할 수 있다는 의미도 가지고 있음을 미리 밝혀둔다. 특히, 디지털 증명서에 이미지 형태가 포함되는 경우(발급기관을 나타내는 로고나 도장 이미지 등) 이 이미지에 정보 텍스트를 삽입하고, 디지털 증명서에 이미지 형태가 포함되지 않을 경우에는 투명 이미지를 생성하여 투명 이미지에 정보 텍스트를 삽입한다. 그리고 정보 텍스트가 삽입된 이미지를 증명서에 넣는 것이다. 물론 육안으로는 아무런 변화가 없다. 즉 삽입되는 정보 텍스트는 증명서가 포함하고 있는 내용에 근거하여 시각적인 확인이 불가능하도록 증명서의 이미지 또는 생성한 투명 이미지에 삽입되며 증명서 발급 후, 증명서의 인증에 사용된다. 본 발명은 증명서에 포함되는 이미지에 정보 텍스트를 삽입하는 방법 및 장치에 관한 것으로서, 삽입된 텍스트는 외부적인 조작이 가해지는 경우, 즉 예를 들면 압축, 필터링(filtering), 리샘플링(re-sampling), 크로핑(cropping) 등 디지털 데이터 처리를 진행하는 경우, 쉽게 제거되며 상기 정보 텍스트의 유무에 근거하여 증명서의 인증을 판단한다. 정보 텍스트의 내용을 제 3 자가 어떤 목적을 가지고 의도적으로 변경할 경우, 삽입되어 있는 정보 텍스트를 추출하여 그 추출 여부 및 처음에 삽입했던 텍스트의 내용과 비교, 그 일치 여부를 판단함으로써 변경사항을 알아낸다.

현재, 인터넷 사용자의 급격한 증가에 따라, 전자 상거래(electric commerce)를 통한 제품의 교역이 활발히 일어나게 되었다. 제품의 교역이나, 온라인/오프라인 거래 과정에는 디지털 증명서의 인증이 없어서는 안될 중요한 절차로 떠오르고 있다. 이러한 전자 상거래 상의 증명서들 뿐 아니라 민원과 관련된 다양한 증명서, 각 학교에서 발급하는 증명서 등 증명서 발급 과정도 전산화, 전자화되는 추세이다. 다만 디지털 문서로 증명서를 발급해 주는 것이 현 시점에서 불가능한 이유는 위조 및 변조의 문제가 가장 큰 걸림돌이라 할 수 있겠다.

디지털 증명서의 인증을 위한 기술에는 인증함수의 적용, 공백제어기법, 바 코드기법 등이 있다. 인증함수를 적용하는 방법은 증명서의 발급자와 소유자 쌍방이 공통의 비밀 값 S를 공유한다고 가정한다. 인증함수에는 비밀 키 알고리즘, 메시지 다이제스트 등 여러 가지가 있는데 대표적인 함수는 해쉬함수(hash Function)이다. 해쉬함수를 이용한 인증은 다음과 같은 절차로 진행된다. 우선, 인증서 발급자는 증명서, 비밀키 값 S, Hash function에 의하여 해쉬 값을 계산하고 증명서에 해쉬 값 결과를 추가한 다음 증명서를 발급 요청자에게 보낸다. 증명서를 발급 받은 자는 자신이 알고있는 세션키 S와 해쉬함수에 의하여 해쉬 값을 구한다. 이렇게 구한 해쉬 값을 인증서 발급자가 보낸 해쉬 값과 비교한다. 같은 경우 인증을 확인하고 다른 경우 인증서가 변조된 것으로 취급한다. 해쉬 값의 비교에 의하여 거래 상대방에 대한 신원확인(인증)과 증명서의 위조, 변조 여부를 검출한다.

이러한 인증함수에 의한 인증은 다음과 같은 약점이 있다.

첫째, 인증은 텍스트문서를 기반으로 처리된다. 만약, 증명서에 이미지나 음성마크와 같은 다른 형태의 문서가 삽입되었을 경우에는 따로 처리해야 되거나, 처리가 불가능하게 된다.

둘째, 문서 데이터가 변경되었는지는 정확히 판단할 수 있지만, 변경된 경우 오리지널 문서의 복원은 할 수 없다.

셋째, 인증서의 무결성을 검증하기 위하여 서명을 추가해야 되므로, 원 증명서의 사이즈가 늘어나게 된다.

공백제어기법이나 글자형 제어기법을 이용한 인증방법은 아스키 코드(ASCII code)를 기반으로 한다. ASCII는 94개의 그림문자와 34개의 여러 종류의 제어용으로만 사용되는 출력되지 않는 문자를 가지고 있다. 그림문자는 26개의 대문자, 26개의 소문자, 10개의 숫자와 32개의 특수문자(%, *, $ 등)를 포함한다. 34개 제어문자는 미리 정해진 규격대로 데이터의 목적지 제어를 하고 출력되어질 텍스트를 배열하는 데 사용한다. 제어문자는 아래와 같은 세가지 유형의 제어문자로 구분이 된다. 규정자(format effectors), 정보 구분자(information separators) 그리고 통신제어문자(communication-control character)가 그것이다. 규정자는 출력의 규정(layout)을 제어하며, backspace(BS), horizontal tabulation(HT), carriage return(CR) 등의 친숙한 타이프용 제어들이 있다. 정보 구분자는 데이터들을 문단 혹은 페이지 등으로 나누는데 사용되며, record separator(RS), file separator(FS) 등이 있다. 통신제어문자는 STX(start of text), ETX(end of Text) 등 전화선을 통해 텍스트 메시지가 전달되는 규정을 만들 때 쓰인다. 대부분의 컴퓨터는 바이트(byte)라는 단위를 8비트 양으로 다룬다. 따라서, ASCII 문자는 대부분 바이트마다 하나씩 저장된다. 여분의 한 비트는 경우에 따라 다른 목적으로

쓰인다. 예로서, 어떤 프린터는 최상위 유효비트를 0으로 맞추고 ASCII를 8비트로 구분한다. 공백제어기법이나 글자형 제어기법은 이러한 제어문자를 이용하여 원 증명서의 텍스트 문서의 여백을 제어하여 삽입하려는 정보를 표현하거나 글자형을 변형하여 삽입정보를 표현하는 기법이다. 이러한 문서의 코드를 대상으로 사용자의 정보를 삽입할 경우 다음과 같은 약점이 있다.

첫째, 실지 증명서의 텍스트 문서 자체가 변하게 된다. 그 변화 자체가 문서에 보이지 않을 따름이다. 따라서 변화된 후 다시 복원 할 경우 복원 신호에 일정한 손실이 있다. 즉, 텍스트 문서의 완전한 복원은 불가능하게 된다.

둘째, 사용자가 정보를 삽입할 수 있는 영역이 적다. 사용자가 원하는 정보를 제한적으로 삽입할 수 밖에 없다.

바코드(Bar code)기법을 적용한 인증은 현재 활발히 적용되고 있다. 바코드는 다양한 폭을 가진 검은 바(Black Bar)와 흰 바(White Bar)의 배열의 패턴으로 정보를 표현하는 부호 체계이다. 바 코드의 인용으로 기존의 사무 처리를 효과적이며 신속, 정확하게 처리할 수 있었다. 국내에서도 그 효용 가치가 인정되면서 유통업을 중심으로 사무자동화, 공장 자동화 등의 다양한 업무 분야에 이용되고 있다. 현재, 각종 ID카드에도 바코드가 이용되고 있다. 학생증의 경우, 바 코드를 넣음으로 해서 도서관 등의 출입 시 카드 인식기에 카드를 인식시키고 인식장치는 바코드를 인식함으로써 신분증의 진위 및 신원정보를 확인하고 있다. 그런데 이러한 바 코드를 이용한 인증은 다음과 같은 약점이 있다.

첫째, 바코드는 다양한 폭을 가진 검은 바(Black Bar)와 흰 바(White Bar)의 배열의 패턴으로 정보를 표현하는 부호 체계이며 따라서 눈으로 식별 가능한 것으로 변조가 가능하다.

둘째, 바코드가 잘 읽히지 않아 스캐너를 여러 번 접촉시키다가 결국에는 키보드로 숫자를 입력하는 경우가 존재한다.

셋째, 바코드가 불명확하거나 유통 과정에서 손상되면 스캐너는 다른 숫자로 읽을 수도 있다.

발명이 이루고자 하는 기술적 과제

종래의 기술들은 증명서에 대한 인증을 함에 있어서 상기 진술한 바와 같이 여러 가지 문제점 및 약점을 지니고 있다. 이러한 문제점으로 인해 증명서가 위조되거나 변조되어, 결국 불법적인 목적으로 오용될 수 있다. 따라서 본 발명의 목적은 각종 온라인/오프라인으로 거래되는 디지털 증명서(신원인증서, 입금확인서, 거래증명서, 졸업증명서, 학위증명서, 현금거래인증서, 상품 영수증, 신용카드 인증서 등)의 인증을 더욱더 확실하게 제공하기 위하여 기존의 방법과는 다른 방법으로 증명서의 인증을 제공할 수 있는, 어떤 정보를 담은 텍스트의 삽입 및 추출 방법과 장치를 제공함에 있다.

본 발명의 또 다른 목적은 필요한 정보의 숨김(Information Hiding)이다. 송신측에서 중요한 정보를 수신측에 보낼 때, 보내려는 정보 텍스트를 이미지 형태의 데이터에 삽입하거나, 이 텍스트를 이미지 형태로 제작한 후 전송하고 수신측에서는 원 정보를 정확하게 복원할 수 있는 방법 및 장치로도 사용될 수 있다. 이러한 과정을 거치면 네트워크 상에서 보편화된 네트워크 탐지기나, 기타 정보 탐지기구에 아무런 주의도 일으키지 않고 전송될 수 있다. 송수신자 사이의 암호화 과정 없이도 정보가 안전하게 보장될 수 있다.

발명의 구성 및 작용

상기 목적을 달성하기 위한 본 발명이 제공하는 방법은 기존의 증명서가 가지는 특성을 분석하고 그 특성에 맞는 정보 텍스트 삽입 기법을 적용하여 증명서를 인증하는 것이다. 본 발명에 따른 디지털 증명서의 인증을 위한 텍스트 삽입 기법은 아래와 같은 단계로 이루어진다. 증명서 분석에 의한 증명서 분류 단계, 분류된 증명서에 따른 정보 텍스트의 생성 단계, 생성된 텍스트를 비트열화 하는 단계, 증명서에 들어 있는 이미지 데이터를 1차원 데이터로 변화시키는 단계, 상기 비트열화 된 텍스트 데이터를 상기 1차원 이미지 데이터에 삽입하는 단계로 구성된다. 삽입된 텍스트의 추출단계는 텍스트가 삽입된 이미지를 선택하는 단계, 삽입된 비트열화 데이터를 추출하는 단계, 추출된 비트열 데이터에서 삽입된 원 텍스트 정보를 판독하는 단계로 구성된다. 증명서 인증을 위한 단계에는 정보 텍스트가 삽입된 증명서를 발급하기 위한 장치 및 소프트웨어, 텍스트가 삽입된 증명서를 인식하기 위한 장치 및 소프트웨어가 포함된다. 본 발명에 따른 장치는 컴퓨터로서 상기 단계에 상응하는 프로그램을 저장한 기록매체로부터 상기 단계들에 해당되는 명령들을 수신하여 텍스트 삽입과 추출을 수행한다.

일반적으로 발행되는 증명서는 아래와 같은 두 가지 특징이 있다.

첫째, 발행되는 증명서에는 증명서를 발행한 법인이나 저작권자를 나타내는, 혹은 발행자임을 확인하기 위한 이미지 형태의 사진이나 인감, 회사의 로고 등이 삽입된다. 이런 이미지의 삽입을 통하여 증명서의 공신력을 높였고, 증명서에 포함되는 개인 정보 및 증명서 양식 등의 디지털화 그리고 발급 절차의 전산화함에 의하여 원격지로의 공문서 발행이 가능하게 되었다.

둘째, 발급되는 증명서는 발급자를 승인하거나 저작권자의 권리를 보호하기 위하여 또는 발급기관의 고유성을 나타내기 위해 지정된 특별용지를 사용하는 경우가 많다. 특별용지에는 발급 기관을 나타내는 이미지가 포함되어 있는 경우가 많고 이미지 형태로 제작되거나 기타 문서 파일 형태를 취한다.

본 발명에서는 이러한 디지털 증명서의 특성에 근거하여, 증명서의 인증을 하기 위한 방법으로 정보를 담은 텍스트의 삽입 기법을 사용한다. 이러한 텍스트의 삽입 기술 및 방법은 디지털 증명서의 보호 및 인증을 위한 것으로서 기존의 증명서 인증 방법과는 차별되는 새로운 기술이다. 특히 본 발명에서 사용하는 텍스트 문서 삽입 방법은 비트열화 된 텍스트 데이터를 삽입하는 방식으로서 발명의 세부적 알고리즘을 포함한 방법과 장치 등은 이하 도면과 함께 설명하기로 한다.

참고로 본 발명의 내용을 설명하기에 앞서 먼저 증명서에 포함되어 있는 이미지들을 살펴볼 필요가 있다. 이는 증명서에 포함되는 이미지의 종류를 분석하고 상기 이미지 특성에 맞는 정보의 삽입을 위한 것이다. 증명서에 들어 있는 이미지 데이터에는 발급처를 나타내는 상징적 이미지나 발급 기관을 나타내는 인감, 도장 이미지 등이 있다. 이러한 이미지 데이터의 실례를 도 14에서 나타내었다.

도 1은 본 발명의 전체적 흐름에 있어 증명서(10)의 인증을 위해 증명서에 텍스트를 임베딩(Embedding)하는 단계에 대한 종합적인 흐름도이다. 도 1의 단계 중 증명서 분류단계에서는 증명서에 이미지 형태의 데이터가 존재하는가(17)의 여부와 증명서는 영문인가(11)의 라는 두 가지 전제 조건에 의하여 증명서를 분류하는데, 구체적인 증명서 분류 과정은 도 4에서 설명하기로 한다. 삽입할 텍스트의 생성 단계(14)에서는 언어에 의한 증명서의 분류 단계를 진행한 후, 영문 기반의 증명서일 경우 텍스트를 삽입하는 방법의 선택 단계(13)를 거치고, 기타 언어 기반의 증명서(12)일 경우 상응한 언어에 의한 텍스트 삽입 방법 선택 과정을 거친다. 영문 기반일 경우 방법 선택 단계에 대한 상세한 내용은 도 5에서 설명하기로 한다. 기타 언어일 경우 언어 선택에 의한 과정은 도 6에서 상세히 설명한다. 생성된 텍스트는 비트열화의 단계(15)를 거쳐서 각각 6, 7, 8, 16 비트 단위로 비트열화 된다. 증명서의 분류단계에서 증명서에 이미지 형태의 데이터가 존재할 경우는 직접 이미지 데이터를 1차원 데이터로 변화시키는 단계(17)를 거친다. 증명서에 이미지 형태의 데이터가 없을 경우는 투명 이미지의 생성단계(18)를 거친다. 투명 이미지의 생성단계는 도 8에서 상세히 설명하기로 한다. 생성된 투명 이미지는 역시 투명 이미지 데이터를 1차원 데이터(20)로 변화시키는 단계(19)를 거친다. 이미지를 1차원 데이터로 변화시키는 단계에서는 RSI기법을 적용하는데 상세한 과정은 도 7에서 설명한다. 비트열화 된 데이터를 1차원 이미지 데이터에 삽입하는 단계(21)는 수학식 4에 의하여 진행된다. 삽입하는 단계에 정보의 확실한 보호를 위하여 키를 사용하게 된다.

도 2는 본 발명의 증명서의 인증을 위한 단계인 정보 텍스트 추출 단계에 대한 종합적인 흐름도이다. 텍스트 추출 단계는 삽입된 텍스트의 추출 단계로서 텍스트가 삽입된 이미지의 선택단계(32), 비트 열 추출단계(33), 증명서의 언어 기반에 의하여 추출된 비트열로부터 텍스트를 구성하는 단계로 이루어 진다. 이미지의 선택 단계에서는 텍스트가 삽입된 이미지를 선택한다. 삽입된 이미지에는 투명 이미지도 포함된다. 선택된 이미지로부터 처음 텍스트 삽입 시 사용한 키 값을 적용하여 비트열 데이터의 추출을 진행하는데, 추출하는 과정은 삽입하는 과정과 마찬가지로 수학식 4에 의하여 이미지의 픽셀 값 정보에서 추출한다. 추출한 비트열 데이터는 영문 증명서(34)인지, 기타 언어의 증명서(38)인지를 구별하여 텍스트 삽입 과정에서 선택한 방법(35 ~ 37, 39 ~ 41) 선택 (도 5와 도 6)의 역 과정을 거쳐서 텍스트 데이터를 생성(42)한다. 생성된 텍스트 데이터가 본 발명에서 추출하려는 데이터이다.

도 3은 디지털 증명서에 대한 특성을 분석한 것이다. 디지털 증명서(50,51)는 도 3에서 보는 바와 같이 특정된 형식의 포맷을 가지며, 흔히 이미지 형태의 데이터(52)가 증명서 내에 존재한다. 본 발명에서는 증명서 자체가 이미지로 된 경우, 증명서에 이미지 형태의 데이터가 존재하는 경우, 증명서에 이미지 데이터가 존재하지 않는 경우에 대하여 모두 가능하다는 것을 알 수 있다. 왜냐하면 증명서가 이미지 형태의 파일일 경우에는 바로 그 증명서 이미지 자체에 정보 텍스트를 삽입하고, 이미지 데이터가 존재하는 경우에는 그 이미지 데이터에 정보 텍스트를 삽입하며, 이미지 데이터가 존재하지 않는 경우에는 투명 이미지를 생성하여 그 이미지에 정보 텍스트를 삽입한 후 증명서에 넣을 수 있기 때문이다.

도 4 는 디지털 증명서를 분류하는 과정을 도면화한 것으로서 분류는 증명서에 사용되는 언어 종류와 증명서의 패턴을 근거로 진행된다. 증명서를 분류하는 단계에서는 다음과 같은 두 가지 조건으로 증명서를 분류한다.

하나는, 디지털 증명서(61)에 이미지 형식의 데이터가 포함되는지를 판단(62)하는 것이고 다른 하나는 증명서가 영문으로 작성된 것인지, 혹은 다른 언어로 작성된 것인지를 판단(63,66)하는 것이다. 상기 두 가지 조건에 근거하여 증명서를 분류한다. 만약 이미지 형식 파일이 존재할 경우, 그 디지털 증명서가 영문으로 작성된 증명서인지, 혹은 다른 언어로 작성된 증명서인지를 판단(63)한다. "영문으로 작성되었는가?" 에 관한 판단은 사용된 ASCII code에 의하여 구분이 가능하고, 이미지는 텍스트와 구별되어 문서 내에서 객체형태(HTML일 경우 링크형태)로 존재한다. 객체에 대한 추적에 근거하여 자동적인 판단이 가능하다. 증명서 전체가 이미지로 되었을 경우, 신경망 등과 같은 문자인식 시스템을 통하여 기존의 증명서에 포함되어 있는 텍스트를 추출하고 사용된 ASCII code의 판단에 근거하여 자동적으로 분류할 수 있다. 증명서 전체가 이미지로 되었을 경우 "이미지포함?"에서는 자동적으로 이미지 포함한것으로 판별된다. 만약 영문으로 작성된 것이면 도 4의 삽입 텍스트 생성부 *1(a)*(64)에 근거하여 삽입할 텍스트 문서가 작성된다. 만약 영문으로 작성된 증명서가 아니라면 도 4의 삽입문서 생성부 *2(a)*(65)에 근거하여 삽입할 텍스트 문서를 생성한다. 이미지 형식의 데이터가 존재하지 않을 경우는 이미지 형식의 데이터가 존재하는 경우보다 투명 이미지를 만드는 단계가 추가된다. 투명 이미지에 대한 생성과정은 도 8에서 설명하기로 한다. 투명 이미지를 작성한 후, 영문으로 작성된 것인지, 혹은 다른 언어로 작성된 증명서인지인지를 판단한다. 만약 영문으로 작성된 증명서라면 도 4의 삽입문서 생성부 *1(b)*(68)에 근거하여 삽입할 텍스트 문서가 작성된다. 만약 영문으로 작성된 증명서가 아니라면 도 4의 삽입문서 생성부 *2(b)*(67)에 의하여 삽입할 텍스트 문서가 생성된다. 여기서 영문인가 아닌가를 판단하는 절차는 아주 중요하다. 본 발명에서는 영문 증명서 외에 한국어, 일본어, 중국어로 된 디지털 증명서를 인증하는 방법을 제공한다. 그리고 다른 어떠한 언어로 된 증명서에도 적용할 수 있음은 물론이다.

도 5는 영문 텍스트 문서를 기반으로 한 증명서에 있어 삽입할 텍스트의 비트열화 과정을 나타낸 것이다. 대상은 도 4에서 증명서의 분류 과정을 거쳐 삽입할 텍스트를 생성하는 단계에 있어, 삽입 텍스트 생성부 *1(a)*(70)와 삽입 텍스트 생성부 *1(b)*(71)에서 생성된 텍스트는 영문으로 된 텍스트 데이터이다. (영문 증명서의 경우 대개는 발급처가 영문을 주 언어로 사용할 것이라 추측되며 따라서 삽입할 텍스트도 영문을 작성되는 것으로 가정한다) 영문으로 된 증명서에 삽입할 정보 텍스트, 즉 8 비트 ASCII code 기반의 텍스트는 다음과 같은 과정을 거쳐서 비트열화 된다. 먼저 기본적으로 비트열화에는 주로 세 가지 방법(72 ~ 74)이 존재하는데, 비트열화 기법의 분류는 주로 참조하는 ASCII code의 테이블에 기초하여 나뉘어 진다. ASCII code는 표 1에 나타나 있다.

[표 1]

	$b_7b_6b_5$							
$b_4b_3b_2b_1$	000	001	010	011	100	101	110	111
0000	NUL	DLE	SP	0	@	P	`	p
0001	SOH	DC1	!	1	A	Q	a	q
0010	STX	DC2	“	2	B	R	b	r
0011	ETX	DC3	#	3	C	S	c	s
0100	EOT	DC4	$	4	D	T	d	t
0101	ENQ	NAK	%	5	E	U	e	u
0110	ACK	SYN	&	6	F	V	f	v
0111	BEL	ETB	‘	7	G	W	g	w
1000	BS	CAN	(	8	H	X	h	x
1001	HT	EM	)	9	I	Y	I	y
1010	LF	SUB	*	:	J	Z	j	z
1011	VT	ESC	+	;	K	[	k	{
1100	FF	FS	,	<	L	\	l	|
1101	CR	GS	-	=	M	]	m	}
1110	SO	RS	.	>	N	^	n	~
1111	SI	US	/	?	O	_	o	DEL

도 5의 방법 1은 표 1에서와 같이 텍스트 문서를 표 1에 기초하여 7 비트로 비트열화(75) 한다. 7비트로 이진화 할 경우, 영문자 100자은 700비트의 비트 열을 구성하게 된다.

영문 코드는 보통 1 바이트(byte) 단위로 처리된다. 나머지 한 비트는 용도에 따라서 조금씩 다르다. 방법 2는 표 1의 7 비트에 나머지 짝수(Even) 혹은 홀수(odd) 형식의 패러티 비트를 추가하여 1 바이트 형태의 코드를 구성하는 것으로 비트열화(76) 한다. 증명서의 삽입할 텍스트 데이터를 비트열화 할 때, 7 비트에 짝수(Even) 혹은 홀수(odd) 형태의 패러티 비트를 추가하여 삽입함을 나타내며 삽입 시, 아스키 짝수 부호표 (ASCII Even code Table) 이거나 아스키 홀수 부호표 (ASCII odd code Table)를 참조함을 말한다. 방법2의 경우, 영문자 100자 당 800비트의 비트 열을 구성하게 된다. (국제표준 ASCII Even code Table과 ASCII odd code Table을 참조)

방법 3은 표 1의 코드의 일부가 사용되지 않는 특성을 고려하여 표 2에서와 같이 사용하는 코드의 개수를 줄여서 6 비트로 표현하고 6 비트 단위로 비트열화(77) 한다. 방법 3의 경우, 영문자 100자 당 600비트의 비트 열을 구성하게 된다.

[표 2]

CHAR	INDEX	CHAR	INDEX	CHAR	INDEX	CHAR	INDEX
a	0	q	16	6	32	`	48
b	1	r	17	7	33	-	49
c	2	s	18	8	34	.	50
d	3	t	19	9	35	/	51
e	4	u	20	SP	36	:	52
f	5	v	21	!	37	;	53
g	6	w	22	"	38	<	54
h	7	x	23	#	39	=	55
i	8	y	24	$	40	>	56
j	9	z	25	%	41	?	57
k	10	0	26	&	42	@	58
l	11	1	27	'	43	[	59
m	12	2	28	(	44	\	60
n	13	3	29	)	45	^	61
o	14	4	30	*	46	{	62
p	15	5	31	+	47		63

따라서 방법 3은 같은 비트 열일 경우, 기타 방법보다 더 많은 텍스트 문서를 삽입할 수 있는 반면 증명서를 구성하고 있는 텍스트 내용과 똑 같은 텍스트의 삽입은 불가능하고, 방법 1과 방법 2는 증명서의 텍스트 내용과 똑 같은 문서의 삽입이 가능하다. 다시 말하면 방법 3은 코드가 줄어들기 때문에 증명서를 구성하고 있는 텍스트의 글자 모양 등을 나타내는 정보를 그대로 표현할 수가 없다. 글자체를 그대로 나타낼 수 없다는 뜻이다.

도 6은 도 4의 분류에 기초하여 삽입할 텍스트를 비트열화 하는 과정을 도면화 한 것이다. 여기서는 영문 증명서가 아닌 한국어, 중국어, 일본어를 기반으로 한 비트열화 과정을 설명한다. 도 4의 삽입문서 생성부 *2(a)*(80)와 삽입문서 생성부 *2(b)*(81)에서 생성된 삽입할 텍스트 데이터(82)는 영문이 아니므로 기타 유니코드(Unicode) 기반의 2 바이트 코드이다. 따라서 증명서의 언어 형태를 구분하여 증명서에 삽입할 텍스트 데이터를 비트열화 한다.

영문 외 언어로 구성된 증명서 즉, 한국어, 중국어, 일본어 등 2 바이트 코드 기반으로 이루어진 증명서는 모두 16 비트 단위로 비트열화(83) 된다. 1993년 5월 배포된 국제표준 부호계 ISO 10646-1/Unicode 1.1는 유니코드가 ISO 10646과 통합함으로써 기술적으로 말하면 ISO 10646와 Unicode version 1.1은 "거의" 같다. ISO 10646-1의 UCS-2와 Unicode가 하나로 합쳐진 것이다. 현재 ISO 10646 개정판인 Unicode 2.0에서는 버전 1.1에서 존재하던 부분적인 문제를 해결했다. 현재 ISO 10646/unicode는 상용화 되고 있다. 국제 표준 부호계는 한국어, 중국어, 일본어 등 을 모두 16 비트 코드 형태로 지원한다. 본 발명에서는 이러한 국제표준에 기반을 두고, 한국어, 중국어, 일본어 등의 언어로 된 디지털 증명서는 비트열화 할 때, 16 비트 단위로 비트열화 하였다. (국제표준 ISO 10646-1 UCS-2 & ISO 10646 Unicode 2.0 / 3.0 code Table참조)

따라서 영문으로 된 증명서와 영문이 아닌 다른 언어로 된 증명서는 비트열화 시, 그 비트열의 길이가 서로 다르게 된다. 즉 영문인 경우, 증명서에 삽입할 수 있는 텍스트의 양이 유니코드(Unicode)기반의 영문 외 언어로 구성된 증명서보다 2배 정도 더 삽입할 수 있는 것이다.

도 7은 디지털 증명서에 포함되는 이미지 혹은 증명서에 이미지가 포함되어 있지 않을 경우 생성한 투명 이미지를 1차원 데이터로 바꾸어주는 과정을 도면화 한 것이다. 이미지를 1차원 데이터로 바꾸어주는 방법에는 여러 가지가 있다. 본 발명에서는 기존에 상용되고 있는 기법 중, 래스터 주사 이미지(Raster Scan Image)기법을 적용했다. 힐버트 주사 이미지(Hilbert Scan Image)기법, 지 스캔 이미지 (Z Scan Image) 기법 등 도 적용이 가능했다.

래스터 주사 이미지는 2차원 데이터를 1차원 데이터로 변환하는 가장 간단하고 널리 사용되고 있는 방법이다. 이미지의 왼쪽 위를 시점으로 하여, 최상위 행(열)부터 차례로 하행(다음 열)인 이미지의 화소(Pixel) 값을 1차원적으로 재 배열시켜, 1차원 이미지 신호를 작성하는 방법이다. 증명서류에서 뽑아낸 이미지가 흑백 이미지(91)일 경우 래스터 주사 이미지 과정을 거쳐서 2차원 데이터를 1차원 데이터로 변환한다. 만약 증명서류에서 뽑아낸 이미지가 컬러 이미지(90)일 경우는 적, 녹, 청 성분으로 분해한 다음 각 래스터 주사 이미지과정을 거쳐서 1차원 데이터로 변환한 다음 적, 녹, 청성분의 순서대로 결합하여 1차원 데이터로 변환한다.

도 8은 투명 이미지의 생성과정을 도면화 한 것이다. 투명 이미지는 발급되는 디지털 증명서의 바탕 환경이 흰색인 경우를 대상으로 하여 제작된다. 바탕 환경이 흰색이 아닌 경우는 바탕 환경이 취하는 색상 정보에 근거하여 투명 이미지를 생성한다. 투명 이미지 생성 과정은 다음과 같다. 정확한 비트 열 정보를 표시하기 위하여 생성하는 투명 이미지는 적어도 2 비트로 양자화 되어 있어야 한다. 2 비트로 양자화 할 경우 이미지는 4개의 픽셀 값(100)으로 구성된다. 여기서 각 픽셀 값을 0, 1, 2, 3로 사상 시키고, 0에는 픽셀 값 RGB=[255,255,255] (101), 1에는 픽셀 값 RGB=[255,255,254](102)을 대응시킨다. 상기, 0과 1을 이용하여 투명 이미지를 생성한다. 0과 1의 선택에서는 모두 0을 취하는 방법, 모두 1을 취하는 방법, 그리고 0과 1을 무작위로 선택하는 방법(103) 등이 있다. 생성하는 투명 이미지의 크기는 삽입되는 텍스트의크기에 의하여 결정(104)하게 된다. 인증을 위한 충분한 양의 텍스트가 삽입 가능하도록 하기 위하여, 투명 이미지의 픽셀의 개수는 적어도 삽입하려는 비트열의 길이보다 커야 한다. 그리고 이미지가 2차원 데이터이므로 두개의 자연수의 곱으로 이루어지는 크기여야 한다.

도 9A는 텍스트 데이터의 삽입 과정을 도면화 한 것이다.

비트열화 된 데이터를 삽입하는 단계는 다음과 같다. 우선, 증명서에 포함된 이미지가 단색일 경우 이미지의 Pixel값의 최대값 Max(Pixel)을 Max(Pixel)-1로 바꾸어준다(120). 증명 서류에서 뽑아낸 이미지가 컬러일 경우 이미지를 색상에 의하여 분해 처리(116)한 후 적, 녹, 청 성분의 픽셀 값의 최대값 Max(Pixel)을 Max(Pixel)-1로 바꾸어준다(117). 이는 정보 텍스트 삽입 시 픽셀 값이 변화되므로 최대값이 0 ~ 2^m-1 범위를 벗어나지 않게 하기 위한 절차이다. 증명서에 이미지가 없는 경우는 새로 생성한 투명 이미지를 그대로 사용한다. 그리고, 도 5 혹은 도 6에서와 같이 삽입할 텍스트(110, 111)를 비트열화(112,113)하고, 도 7에서와 같이 이미지 데이터를 1차원 데이터로 한다(118). 비트열화 된 텍스트 데이터를 $B(i)$(114)로 표시하였다. 상기 이미지를 1차원 데이터로 변환시킨 결과 데이터를 $image(i)$ 로 표시하고, 1차원 데이터의 정의역 범위를 칼라 이미지, 단색이미지로 분류하여 각각 수학식 1와 수학식 2에 표시하였다. Nx M은 이미지의 가로 세로의 크기를 나타낸다.

수학식 1

$$0 \le i \le N \times M$$

수학식 2

$$0 \le i \le N \times M \times 3$$

수학식 3은 1차원 데이터가 취할 수 있는 값의 범위이다. 여기서 m 은 그 1차원 데이터의 양자화 비트 수를 나타낸다. Key_{SN}는 비트열화 된 데이터를 1차원 데이터로 변환시킨 상기 이미지 데이터에 삽입할 때, 첫 삽입위치를 지정해주는 키 값을 나타내고 $image(key_{SN} + 1)$는 1차원 데이터에서 실제 삽입되는 위치를 나타낸다.

수학식 3

$$0 \le image(i) \le 2^m - 1$$

삽입의 원리(122)는 다음과 같다.

비트열화 된 값과 삽입 위치의 값을 비교하여 수학식 4를 만족할 경우, 이미지의 픽셀 값인 수학식 5는 변화시키지 않는다. 만약 수학식 4을 만족시키지 않으면, 1차원화 된 이미지의 픽셀 값을 $image(key_{SN}+i)+1$에 의하여 변환시킨다.

수학식 4

$$image(key_{SN}+i) \equiv B(i) \mod 2$$

수학식 5

$$image(key_{SN}+i)$$

구체적 예로 도면 9B에 도시된 바와 같이, 6 X 6 픽셀 이미지(125)에 텍스트 "four-life" (126)를 삽입하는 과정을 방법3에 근거하여 진행하면 다음과 같다(즉, 표2를 참조). 먼저 6 X 6 이미지를 RSI 기법을 적용하여 1차원화된 데이터 *Image(i)*(127)로 만든다. 표2를 보면 "f"는 5, "o"는 "14"에 대응된다. 따라서 "four-life"를 표 2의 인덱스에 근거하여 참조하여 십진수 열 " 5, 14, 20, 17, 49, 11, 8, 5, 4"를 얻는다. 이를 6 비트를 단위로 하는 비트 열로 바꾸어서 비트열화 된 데이터 *B(i)*를 얻는다. 수학식 4에 근거하여 image(1)=123과 B(1)=0는 이므로 123+ 1을 적용하여 픽셀 값을 124로 바꾸어 준다. 바꾸어 말하면, 일차원화된 이미지의 첫번째 픽셀의 값이 123 임으로 이를 2 로 나누면 나머지가 1 이 된다. 그 나머지 1 은 비트열의 첫번째 비트 값 0 과 동일하지 않음으로 첫번째 픽셀의 원래의 값에 1을 더하여 124로 만든다. 반면 두번째 픽셀 Image(2)=124는 2로 나눈 나머지가, 비트열의 두번째 비트값인 0 와 동일함으로 그 픽셀 값을 변화시키지 않는다. 같은 방법으로 모든 B(i)의 비트 열 정보를 6 X 6 이미지에 삽입한다. 삽입 후 다시 역 RSI 기법(123)을 적용하여 이미지를 구성(124)한다. 이렇게 구성한 이미지는 텍스트가 삽입된 이미지(129)이다.

키는 증명서 소유자 혹은 증명서 발급자가 임의로 정한 1~8자리의 십진수에 의하여 정해진다. 물론 이 키를 그대로 사용하는 것은 아니다. 이하 수학식 6, 7, 8에 의한 수학적인 연산을 거쳐서 키의 값이 지정한 범위(이하 수학식 6에서의 D)에 들어가도록 한다. 키의 값이 지정한 범위로 바뀌는 과정에 관하여는 증명서 발급자는 신경을 쓸 필요가 없다. 단지 처음에 정한 키의 정보만 기억하고 있으면 된다. 본 발명에서 사용되는 키는 아래와 같은 두 가지 조건의 제약을 받는다.

첫번째 제약 조건은 다음과 같다.

키는 증명서에서 추출한 이미지의 픽셀의 개수 즉, 이미지의 크기와 밀접한 관계가 있으며 이미지 크기의 영향을 받는다. 이미지 크기가 실제 삽입할 수 있는 텍스트에서 생성한 비트열의 범위가 된다. 여기서*image(i)* 의 정의역 *i*가 가지는 최대치를 *Max(image,i)* 라고 한다.

두번째 제약 조건은 다음과 같다. 삽입할 비트열의 길이는 *B(i)*에서 *i* 가 가지는 값의 최대치가 얼마인가에 따라 결정되게 된다. 여기서의 *B(i)* 최대치를 *Max (B,i)*라고 한다. 키 Key_{SN}가 가질 수 있는 값의 범위는 *D*라고 한다. *D*의 값은 비트열의 최대치 *Max(B,i)* 와 이미지 사이즈에 얻은 삽입할 수 있는 비트열의 최대치 *Max(image,i)*와 다음과 같은 관계가 있다.

수학식 6

$$D = \{\forall key_{SN} \mid 0 \le key_{SN} \le Max(image,i) - Max(B,i)\}$$

만약, 입력한 Key_{SN}의 값이 수학식 7에서 표시한 값보다 크면 키 값을 수학식 8와 같은 연산을 경과한 후 사용한다. 입력한 Key_{SN}의 값이 수학식 7에서 표시한 값보다 작으면 그대로 사용한다. 실제 사용되는 키 값을key_{SN1} 라고 하면 그 값은 수학식 8와 같다.

수학식 7

$$Max(image,i) - Max(B,i)$$

수학식 8

$$key_{SN1} = Key_{SN} \mod Max(image,i) - Max(B,i)$$

키는 아래와 같은 용도로 사용된다.

첫째, 비트열화 된 텍스트 문서의 데이터를 1차원화 된 이미지 데이터로의 삽입 위치를 지정하는 용도로 사용된다.

둘째, 증명서의 발급자, 소유자가 자신만의 정보임을 확인할 때 사용된다. 이는 다음과 같은 공격을 막기 위함이다. 키 값을 이용하여 어떤 사람이 텍스트 문서를 위조한 후, 증명서의 이미지에 삽입된 문서도 본 알고리즘을 이용하여 확인하려고 하는 경우를 막을 수 있다.

셋째, 키 값을 이용하여 삽입 구간을 정했기에 만약 키 값의 보안을 유지하면 삽입한 텍스트 데이터 자체의 안정성도 보장되게 된다.

삽입 알고리즘에서 볼 수 있는 바와 같이 증명서에 의거하여 이미지를 선택 또는 생성하고, 증명서에 근거하여(필요한 정보/전체 증명서 내용 등) 삽입할 정보 텍스트를 생성한다. 증명서에서 선택한 이미지에 증명서에 의거하여 작성한 텍스트 정보를 삽입하게 되면, 원본 이미지와 시각적으로 구분할 수 없는 새로운 이미지가 작성된다. 이러한 과정을 거쳐서 텍스트의 삽입을 실현하여 증명서의 인증을 할 수 있다.

도 10은 텍스트 데이터의 추출 과정을 도면화 한 것이다. 추출 알고리즘은 다음과 같은 절차로 진행된다. 우선, 증명서(130)에 포함되어 있는 개인이나 발급자의 인감이미지, 사용자의 증명사진, 또는 발행회사의 로고와 이미지와 같은 이미지에서 정보 텍스트가 삽입된 이미지(투명 이미지 포함)를 선택(131)한다. 그리고, 선택한 이미지를 래스터 주사 이미지 기법(133)을 이용하여 1차원 데이터로 변환(134) 시킨다. 다음, 사용자 혹은 발행자가 가지고 있는 키에 의하여 삽입이 시작된 위치를 찾아내고 1차원 데이터 $image(Key_{SN}+1)$에서 시작하여 픽셀 값의 기우성에 근거하여 $image(key_{SN}+1) \equiv B(i) \bmod 2$ 를 만족시키는 비트열 $B(i)$(135) 를 구성한다. 이어서, 구성한 비트 열(136) 를 영문 기반의 증명서일 경우 삽입 시 적용한 삽입 방법1, 삽입 방법2, 삽입 방법3을 선택하여 각각, 8비트, 7비트, 6비트 단위로 결합하여 십진수로 바꾸어준다(137). 영문 기반의 증명서가 아닐 경우, 16비트 단위로 결합한 후 십진수로 바꾸어준다. 그리고, 구성한 십진수 열을 삽입 방법이 참조한 테이블에 근거하여 텍스트 데이터로 전환한다(138). 전환하여 얻어진 결과가 복원하려는 텍스트 데이터(139)가 된다.

도 10에서 볼 수 있는 바와 같이 삽입된 텍스트를 복원하는 과정은 삽입 과정의 역 과정이다.

본 발명의 구체적인 알고리즘은 기본적으로 세가지 단계로 이루어졌다. 즉, 삽입할 정보 텍스트의 생성단계, 정보 텍스트의 삽입단계, 정보 텍스트의 추출단계가 그것이다. 알고리즘의 성능에 직접적인 영향을 미치는 것은 정보 텍스트의 삽입과정에서의 이미지의 변화와 텍스트로 삽입되는 정보의 양이다. 이에 대하여 보다 상세히 설명하면 다음과 같다. 먼저 이미지 변화에 대해 언급하기로 한다.

텍스트 문서의 삽입과정은 실제 이미지를 변화시키는 과정이라고 할 수 있다. 즉 이미지에 사용자가 지정한 텍스트를 삽입하여 이미지를 변화시키는 것인데, 이는 이미지에 노이즈(Noise)를 삽입하는 과정이라고도 할 수 있다. 삽입한 노이즈의 크기에 의하여 원 이미지와 텍스트가 삽입된 이미지가 구분된다. 본 발명에서 인증서로 발급하여 발급 요청자에게 전달되는 증명서에는 원래 최초로 생성되었던 증명서 안에 포함된 이미지가 들어 있지 않다. 즉 발급된 증명서에 삽입되어 있는 이미지는 발급자가 생성한 정보 텍스트가 삽입된 이미지이다. 그리고 최초의 원 이미지는 정보 텍스트를 삽입한 후 더 이상 필요하지 않다.

도 11, 도 12, 도 13는 텍스트가 삽입되기 전의 단색 이미지(140), 칼라 이미지(150), 인감 이미지(160)와 텍스트가 삽입된 후의 이미지(142,152,162)를 히스토그램(141,143,151,153,161,163)과 함께 비교해본 결과이다. 픽셀 값의 변화가 작기 때문에 시각적으로는 구별할 수 없는 것으로 나타났다. 픽셀 값 변화는 삽입된 영역에서만 진행된다.

도 11의 히스토그램을 비교해보면 텍스트 데이터가 많이 삽입된 적(Red) 성분에서 변화가 생겼음을 알 수 있다. 도 12, 도 13는 칼라 이미지 경우와 인감 이미지에 대한 성능 평가이다. 단색 이미지, 칼라 이미지, 그리고 인감 이미지에 텍스트를 삽입하여 원 이미지와의 변화를 비교 분석한 결과 이미지의 픽셀 값의 변화를 시각적으로 확인하는 것은 불가능함을 알 수 있다.

다음으로 삽입되는 정보의 양에 대해 살펴본다.

일련의 텍스트로 삽입되는 정보의 양은 증명서에 포함되어 있는 이미지의 사이즈와 그 이미지의 양자화 수준과 직접적인 관계가 있다. 아래의 표는 양자화 비트 수와 그 이미지 사이즈에 따른 텍스트 삽입량의 계산 결과이다. 삽입한 정보의 양에서 볼 수 있는 바와 같이 증명서에 이미지가 포함되어있을 경우에는 그 내용의 제한을 거의 받지 않는다. 보통 증명서는 500자~3,000자 정도의 텍스트 내용을 포함한다. 따라서 삽입할 정보 텍스트의 사이즈의 제한을 거의 받지 않는다.

[표 3]

이미지 사이즈	양자화 비트 수	삽입할 수 있는 정보 양
256 by 256	칼라(24bit)	Max 32,768자(6bit단위)
256 by 256	단색(8bit)	Max 10,922자(6bit단위)
153 by 134	칼라(24bit)	Max 10,251자(6bit단위)
153 by 134	단색(8bit)	Max 3,417자(6bit단위)

그러면, 본 발명에서 제안하는 방법을 실시한 이미지를 제 3자가 다양한 공격(조작 등) 및 어떠한 처리를 하였을 때의 영향에 대해 설명하기로 한다.

이미지 처리에는 많은 종류가 있다. 본 발명에서는 텍스트가 삽입된 이미지에 대해 여러 가지 이미지 처리를 실험적으로 진행해 보았다. 텍스트가 삽입된 이미지에 JPEG과 같은 주파수 공간에서의 압축을 진행할 경우 삽입된 텍스트가 사라졌다. 유사한 방법인 필터를 적용했을 경우에도 삽입된 텍스트를 복원할 수 없었다. 다른 이미지 처리기법으로 히스토그램의 균일화(Histogram equalization), 이미지 예리화(Image sharpening), 이미지 크로핑(Image cropping)을 적용해 보았다. 그 결과 여전히 삽입된 텍스트 문서를 복원할 수 없었다. 회전을 비롯한 일부 기하학적인 이미지 처리를 제외한 거의 모든 이미지 처리 기법에 의하여 삽입된 텍스트를 추출할 수 없었다.

아래의 표 4는 이미지 처리에 대한 분석 결과이다. (X-추출 불가능, O-추출 가능)

[표 4]

이미지 / 처리기법	칼라 이미지	단색 이미지	인감 이미지
JPEG압축	X	X	X
Image cropping	X	X	X
Equalization	X	X	X
Sharpening	X	X	X
평활화(특징추출, 미분)필터링	X	X	X
	X	X	X
기하학적 회전변화	O	O	O

즉 본 발명에서 제안한 방법은 "인증"의 목적을 위해 반드시 요구되는 취약한 워터마크(Fragile Watermark)의 성질을 그대로 담고 있다. (Fragile Watermark에 대해서는 본 출원인이 출원한 특허를 참고하기로 한다)

증명서의 변조는 증명서의 내용으로 들어있는 텍스트(이름이나 날짜 등)를 변조하는 경우와 증명서에 포함되어 있는 이미지를 변조하는 경우가 있다. 증명서에 포함되어 있는 텍스트 내용을 변조할 경우, 이미지에 삽입한 증명서 내용 정보 텍스트를 추출함으로써, 그 변조 상황을 알아낼 수 있으며, 원 텍스트 문서를 복원할 수 있다. 즉 상기 결과에 의하여 증명서에 포함되어 있는 텍스트 내용 모두를 증명서에 있는 이미지에 삽입할 수 있으므로 삽입된 텍스트를 추출함으로써 원래 증명서가 가지고 있는 내용을 모두 복원해 낼 수 있으면 위변조 여부도 확인할 수 있는 것이다. 또한 증명서에 포함되어 있는, 텍스트가 삽입된 이미지를 변조할 경우, 삽입된 텍스트는 추출할 수 없게 된다. 상기 이미지에 삽입된 텍스트를 검출할 수 없을 경우, 증명서는 무효로 취급된다.

증명서의 위조는 증명서의 일부를 자신이 원하는 형태로 바꾸는 데서 이루어진다. 만약 증명서를 위조하려 할 경우, 증명서의 발급자나 증명서 이용자가 사용하는 키를 알아야 된다. 키를 모를 경우, 상기 이미지에 텍스트를 삽입할 수는 있어도 인증하는 것은 불가능하게 된다. 따라서 위조 상황도 검출할 수 있다.

도면 15 에서 보이듯, 본 발명의 다른 실시예에 따르면 바코드와 위에서 언급한 텍스트 삽입방법의 결합적인 사용도 가능하다. 이는 바코드를 증명서에 삽입하는 이미지에 추가하여 사용하는 방법이다. 즉 현재 증명서에 들어가는 이미지 형태 중에 바코드(171)를 추가하고 본 발명을 적용하는 방법이다. 이 방법의 형식에는, 바코드와 다른 이미지를 통합하여 한 이미지로 적용하는 경우와 바코드와 텍스트를 삽입하는 이미지를 따로 적용하는 경우, 두 가지로 나누어 볼 수 있다.

통합하여 한 이미지로 적용할 경우, 바코드와 겹쳐지는 이미지 부분에는 블록처리를 진행하여 텍스트를 삽입하지 않음으로써 바코드에 대한 안정성과 텍스트에 대한 안정성을 유지할 수 있다. 따로 사용하게 되는 경우 바코드를 통해서는 증명서의 인증을 쉽고 빠르게 진행하고, 이미지에 삽입되어 있는 텍스트를 추출하는 본 발명의 과정을 한번 더 진행함으로써 이중 인증으로 보안성, 안전성을 높일 수 있다는 장점이 있다. 특히 바코드를 사용하게 될 경우, 증명서를 위조 또는 변조하려는 자가 단순히 바코드만을 보고 그 부분을 공격할 경우, 보이지 않게 텍스트를 삽입한 이미지(170)에 대한 부분은 소홀히 지나칠 수 있으므로 이중 보안의 효과를 지닌다.

도시하지는 않았지만, 위에서 상술한 텍스트 삽입과 추출은 컴퓨터를 사용하여 실행한다. 또한, 통상의 프로그래밍 언어로, 대상 디지털 증명서, 즉, 전자문서를 읽고 텍스트와 이미지를 판별하고 텍스트를 이미지에 삽입하는 프로그램을 작성한 다음 기록매체에 저장하여 컴퓨터에서 실행되게 한다. 통상의 기술수준을 가진 당업자라면 용이하게 본원발명을 구현하는 다양한 프로그램들을 개발할 수 있다. 즉, 널리 알려진 바와 같이, 그 프로그램과 관련 데이터는 플로피 디스크와 같은 외부 기억매체로부터 컴퓨터의 하드 디스크에 일시적으로 저장될 수 있고 상기 텍스트 삽입 프로그램이 동작되게 될 때 디스크로부터 RAM 에 저장될 수 도 있다. 프로그램 코드를 제공하기위한 기억매체로서, 예를 들면 플로피 디스크, 하드 디스크, 광 디스크, 광자기 디스크, CD-ROM, CR-R, 자기 테이프, 불휘발성 메모리 카드, ROM 등이 사용될 수도 있다. 본 발명은 전술된 기억매체에 인가되는 경우, 전술한 삽입단계들에 대응되는 프로그램 코드들은 기억매체에 저장된다.

발명의 효과

본 발명은 각종 디지털 증명서의 진위를 확실하게 인증할 수 있다. 증명서의 위조나 변조를 막음으로써 이로 인해 발생할 수 있는 각종 피해를 막을 수 있는 효과를 갖는다. 본 발명은 증명서의 위조/변조를 가려낼 수 있고, 위조/변조 되었을 경우 위조/변조 된 내용을 복구할 수 있다. 그리고 나아가 증명서 발급자의 서명을 안전하게 수신자에게 전달할 수 있고 삽입되어 있는 텍스트의 내용을 추출해 봄으로써 거래 상대방에 대한 확실한 신원 확인 또한 가능하다.

(57) 청구의 범위

청구항 1.

프로세서, 및 상기 프로세서에 의해 실행되는 프로그램을 저장한 컴퓨터 판독가능 기록매체를 포함하는 전자문서 인증 컴퓨터 시스템에서의 정보 텍스트 삽입 방법에 있어서,

상기 방법은,

상기 전자문서 중 텍스트 데이터의 최소한 일부분을 비트열로 변환하고, 상기 전자문서 중 이미지 데이터를 일차원의 데이터로 변환하는 단계;

상기 일차원화된 데이터의 단위 데이터에 나머지 연산을 적용하여 나머지를 구하고 그 나머지를 상기 비트열의 비트 값과 비교한 결과에 기초하여, 상기 일차원화된 데이터를 변형하여 제 1 잠정 데이터를 발생시키는 단계;

상기 제 1 잠정 데이터를 이미지 형태로 변환하여 제 2 잠정 데이터를 발생시키는 단계; 및

상기 전자문서 중 이미지 데이터를 상기 제 2 잠정 데이터로 치환하는 단계를 포함하며,

상기 프로그램은 상기 프로세서로 하여금 상기 방법을 수행하게 하는, 정보 텍스트를 삽입하는 방법.

청구항 2.

삭제

청구항 3.

삭제

청구항 4.

제 1 항에 있어서,

상기 일차원화된 데이터의 각 단위 데이터를 상기 나머지 연산자를 사용하여 나머지를 구하고 그 나머지를 상기 비트열의 일개의 비트 값과 비교하여 동일하면 상기 단위 데이터를 그대로 유지하고 상이하면 상기 단위 데이터에 1 을 더하여 변형하는, 정보 텍스트를 삽입하는 방법.

청구항 5.

제 1 항에 있어서,

상기 텍스트 형태의 데이터의 언어종류를 판별하고 상응하는 각 언어에 상응하는 비트열로 변환하는 단계를 포함하는, 정보 텍스트를 삽입하는 방법.

청구항 6.

프로세서, 및 상기 프로세서에 의해 실행되는 프로그램을 저장한 컴퓨터 판독가능 기록매체를 포함하는 전자문서 인증 컴퓨터 시스템에서의 정보 텍스트 삽입 방법에 있어서,

상기 방법은,

상기 전자문서 중 텍스트 데이터의 최소한 일부분을 비트열로 변환하고, 상기 전자문서의 바탕을 표시하는 형태의 데이터를 일차원의 데이터로 변환하는 단계;

상기 일차원화된 데이터의 단위 데이터에 나머지 연산을 적용하여 나머지를 구하고 그 나머지를 상기 비트열의 비트 값과 비교한 결과에 기초하여, 상기 일차원화된 데이터를 변형하여 제 1 잠정 데이터를 발생시키는 단계;

상기 제 1 잠정 데이터를 이미지 형태로 변환하여 제 2 잠정 데이터를 발생시키는 단계; 및

상기 전자문서의 바탕을 표시하는 형태의 데이터를 상기 제 2 잠정 데이터로 치환하는 단계를 포함하며,

상기 프로그램은 상기 프로세서로 하여금 상기 방법을 수행하게 하는, 정보 텍스트를 삽입하는 방법.

청구항 7.

제 1 항에 있어서,

상기 텍스트 데이터의 최소한 일부분을 블록암호화 하거나 또는 메시지 다이제스트 값을 구한후 비트열로 변환하는, 정보 텍스트를 삽입하는 방법.

청구항 8.

제 1 항에 있어서,

상기 일차원의 비트들로 변환하는 단계는 래스터 스캔 이미지 (Raster Scan Image), 힐버트 스캔 이미지 (Hilbert Scan Image), 또는 지 스캔 이미지 (Z Scan Image) 기법 중 하나를 선택하는, 정보 텍스트를 삽입하는 방법.

청구항 9.

제 1 항에 있어서,

상기 비트열을 상기 일차원 비트들에 삽입할 때 그 삽입위치는 공개키 또는 비밀키에 근거해서 선택하는, 정보 텍스트를 삽입하는 방법.

청구항 10.

제 1 항에 있어서,

상기 이미지 데이터는 바코드를 포함하는, 정보 텍스트를 삽입하는 방법.

청구항 11.

프로세서, 및 상기 프로세서에 의해 실행되는 프로그램을 저장한 컴퓨터 판독가능 기록매체를 포함하는 전자문서 인증 컴퓨터 시스템에서의 정보 텍스트 추출 방법에 있어서,

상기 방법은,

상기 전자문서에서 텍스트 데이터와 이미지 데이터를 판별하여 이미지 데이터 일부를 선택하는 단계;

상기 선택된 이미지 데이터 일부를 일차원의 데이터로 변환하는 단계;

상기 일차원 데이터의 각 단위 데이터에 나머지 연산을 행한 결과에 기초하여, 일련의 비트열을 발생시키는 단계; 및

상기 일련의 비트열을 상기 텍스트 데이터와 동일한 형태로 변환하는 단계를 포함하며,

상기 프로그램은 상기 프로세서로 하여금 상기 방법을 수행하게 하는, 정보 텍스트를 추출하는 방법.

청구항 12.

제 11 항에 있어서,

상기 일차원의 데이터에 근거하여 일련의 비트열을 발생시키는 단계는 상기 일차원 데이터의 각 단위 데이터를 2로 나누어 나머지를 구하고 그 나머지들로 상기 일련의 비트열을 발생시키는 단계를 포함하는, 정보 텍스트를 추출하는 방법.

청구항 13.

정보 텍스트를 삽입하기 위한 시스템에 있어서,

프로세서 및 상기 프로세서에 의해 실행되는 프로그램을 저장한 컴퓨터 판독가능 기록매체를 포함하며,

상기 프로그램은,

전자문서 중 텍스트 데이터의 최소한 일부분을 비트열로 변환하고, 상기 전자문서 중 이미지 데이터를 일차원의 데이터로 변환하는 단계;

상기 일차원화된 데이터의 단위 데이터에 나머지 연산을 적용하여 나머지를 구하고 그 나머지를 상기 비트열의 비트 값과 비교한 결과에 기초하여, 상기 일차원화된 데이터를 변형하여 제 1 잠정 데이터를 발생시키는 단계;

상기 제 1 잠정 데이터를 이미지 형태로 변환하여 제 2 잠정 데이터를 발생시키는 단계; 및

상기 전자문서 중 이미지 데이터를 상기 제 2 잠정 데이터로 치환하는 단계가 상기 프로세서에서 실행되게 하는, 정보 텍스트 삽입 시스템.

청구항 14.

컴퓨터에,

전자문서 중 텍스트 데이터의 최소한 일부분을 비트열로 변환하고, 상기 전자 문서 중 이미지 데이터를 일차원의 데이터로 변환하는 단계;

상기 일차원화된 데이터의 단위 데이터에 나머지 연산을 적용하여 나머지를 구하고 그 나머지를 상기 비트열의 비트 값과 비교한 결과에 기초하여, 상기 일차원화된 데이터를 변형하여 제 1 잠정 데이터를 발생시키는 단계;

상기 제 1 잠정 데이터를 이미지 형태로 변환하여 제 2 잠정 데이터를 발생시키는 단계; 및

상기 전자문서 중 이미지 데이터를 상기 제 2 잠정 데이터로 치환하는 단계를 실행시키기 위한 프로그램을 기록한, 컴퓨터 판독가능 기록매체.

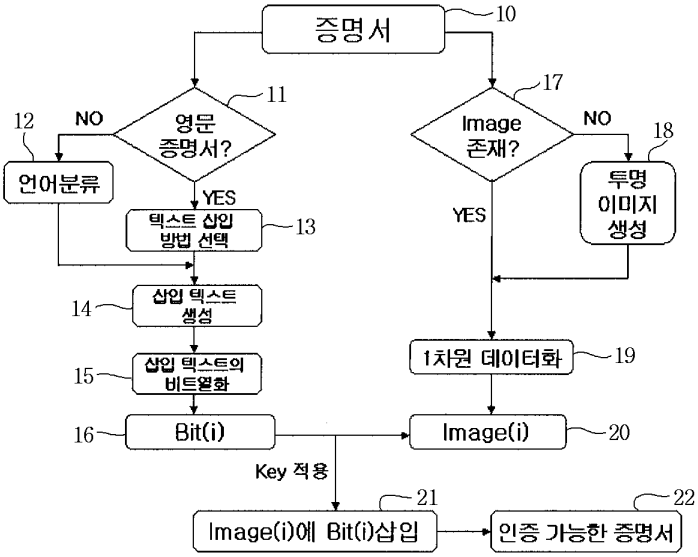
청구항 15.

제 6 항에 있어서,

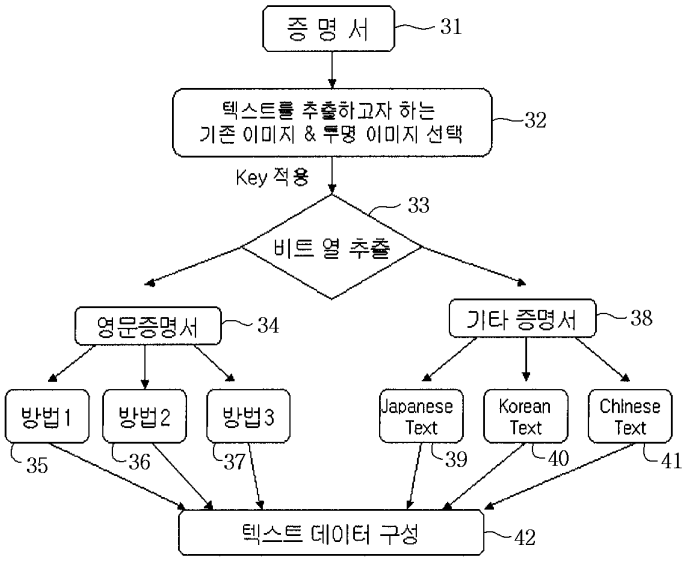
상기 전자문서의 바탕을 표시하는 형태의 데이터는 바코드를 포함하는, 정보 텍스트를 삽입하는 방법.

도면

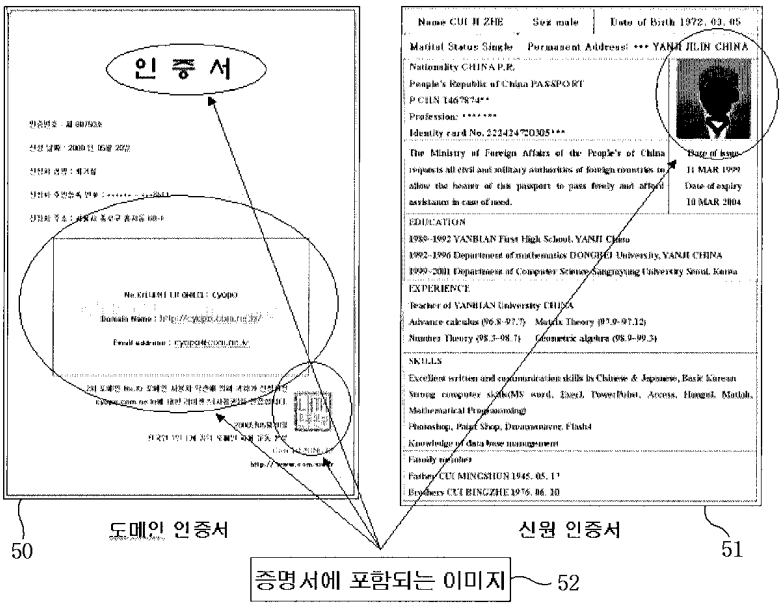
도면1


증명서
10
영문 증명서?
11
NO
12
언어분류
YES
텍스트 삽입 방법 선택
13
삽입 텍스트 생성
14
삽입 텍스트의 비트열화
15
Bit(i)
16
Image 존재?
17
NO
18
투명 이미지 생성
YES
1차원 데이터화
19
Image(i)
20
Key 적용
Image(i)에 Bit(i)삽입
21
인증 가능한 증명서
22

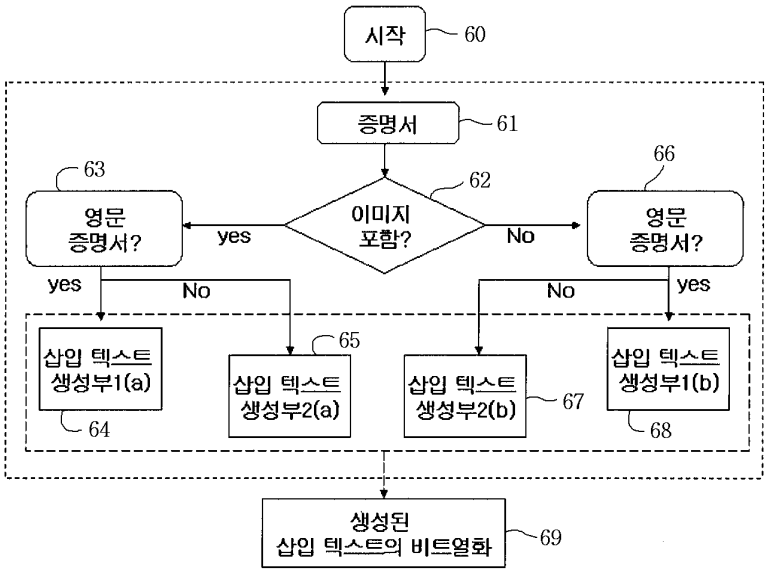
도면2


증 명 서
31
텍스트를 추출하고자 하는 기존 이미지 & 투명 이미지 선택
32
Key 적용
비트 열 추출
33
영문증명서
34
기타 증명서
38
방법1
35
방법2
36
방법3
37
Japanese Text
39
Korean Text
40
Chinese Text
41
텍스트 데이터 구성
42

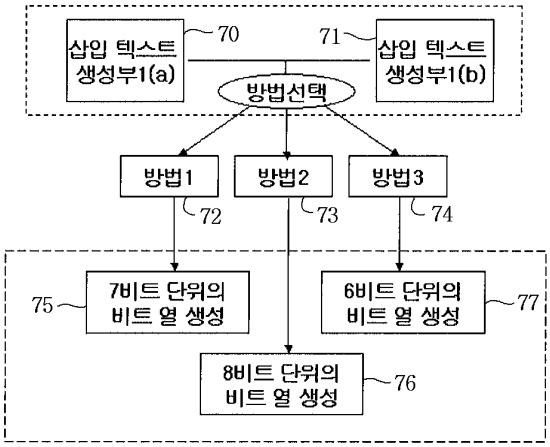
도면3


인 증 서
Name CUI H ZHE Sex male Date of Birth 1972. 03. 05
Martial Status Single Permanent Address: *** YANJI JILIN CHINA
Nationality CHINA P.R.
People's Republic of China PASSPORT
Profession: ******
Identity card No. 222424720305***
The Ministry of Foreign Affairs of the People's of China requests all civil and military authorities of foreign countries to allow the bearer of this passport to pass freely and afford assistance in case of need.
Date of issue 11 MAR 1999
Date of expiry 10 MAR 2004
EDUCATION
1988-1992 YANBIAN First High School, YANJI China
1992-1996 Department of mathematics DONGHEI University, YANJI CHINA
1999-2001 Department of Computer Science Sangmyung University Seoul, Korea
EXPERIENCE
Teacher of YANBIAN University CHINA
SKILLS
Family members
Brothers CUI BINGZHE 1976. 06. 10
도메인 인증서
신원 인증서
50
51
증명서에 포함되는 이미지
52

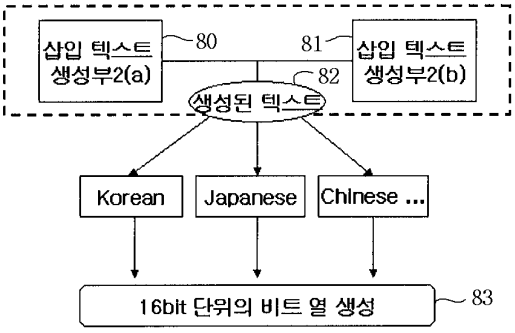
도면4


시작
60
증명서
61
62
이미지 포함?
yes
No
63
영문 증명서?
66
영문 증명서?
yes
No
No
yes
삽입 텍스트 생성부1(a)
64
65
삽입 텍스트 생성부2(a)
삽입 텍스트 생성부2(b)
67
삽입 텍스트 생성부1(b)
68
생성된 삽입 텍스트의 비트열화
69

도면5

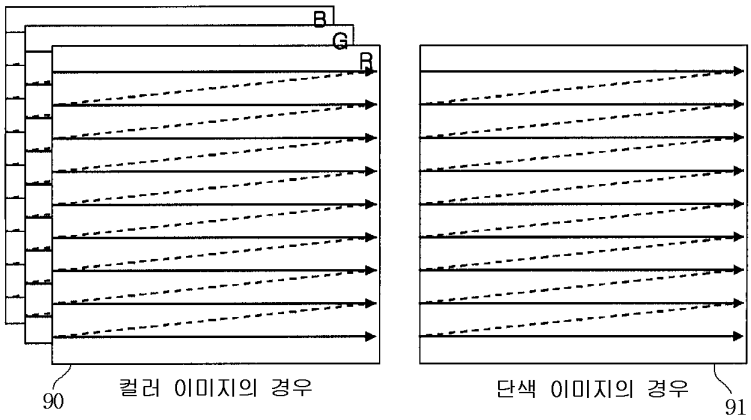


도면6



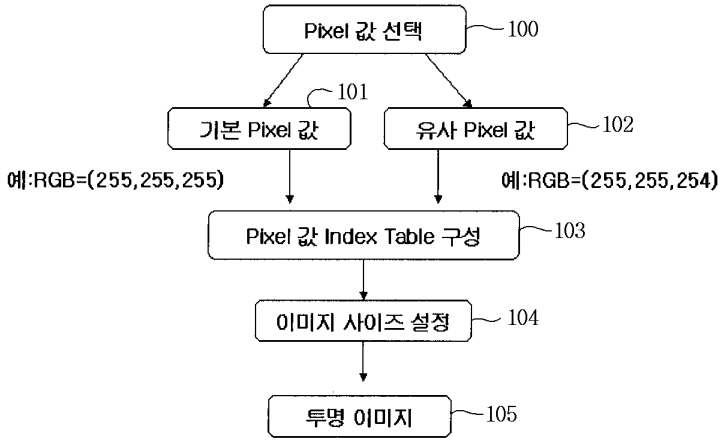
도면7

이미지 데이터의 1차원으로의 변환

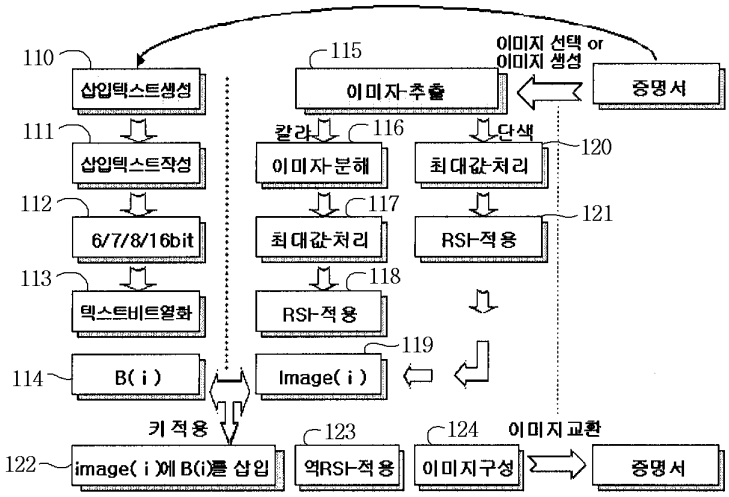


칼라 이미지의 경우 R-성분을 1차원으로 배열한 후
이어서 G-성분과 B-성분을 배열한다

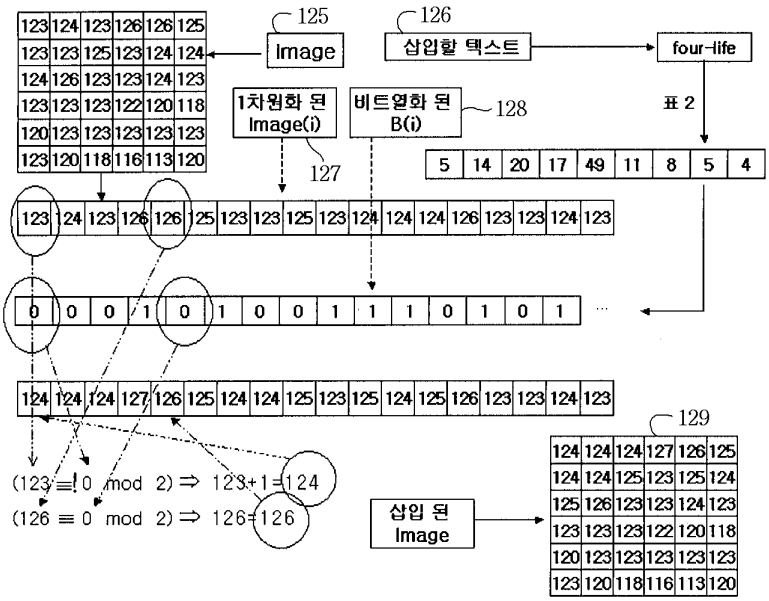
도면8



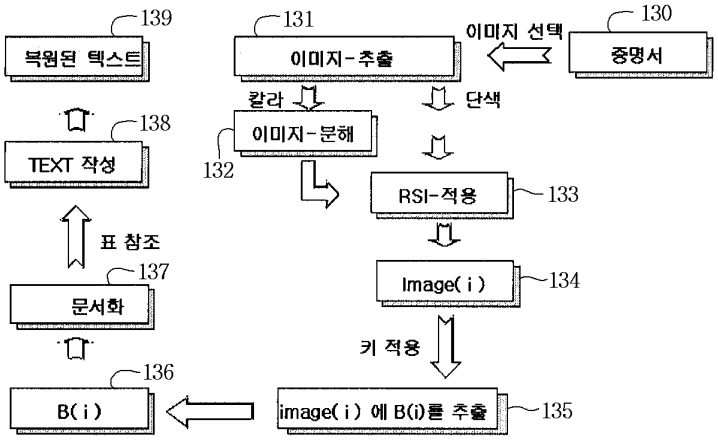
도면9a



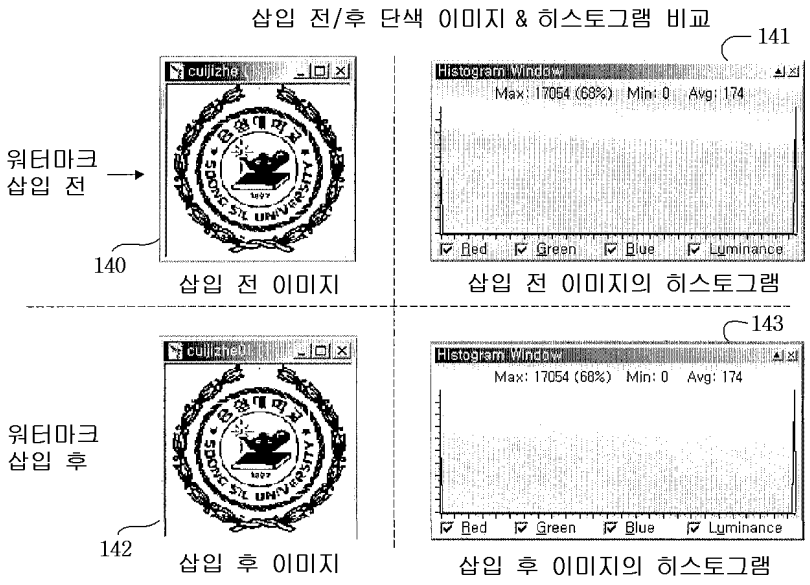
도면9b



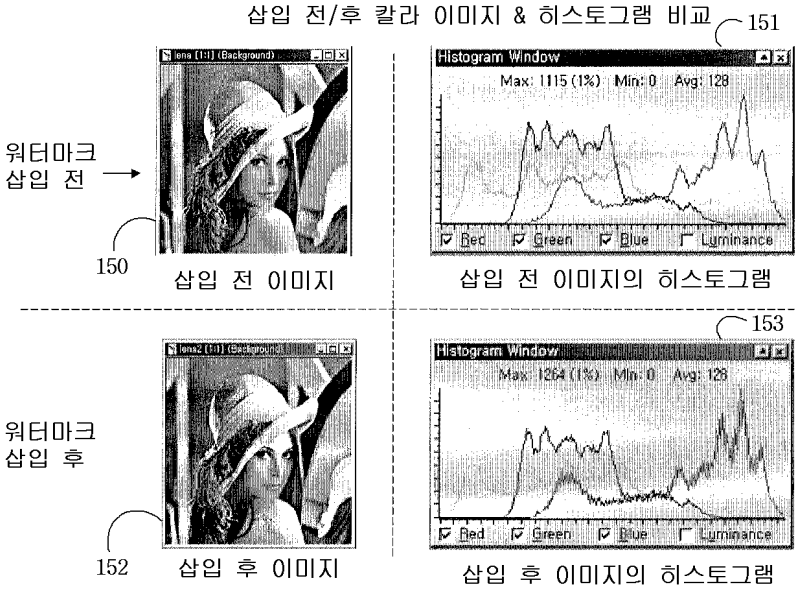
도면10


139
복원된 텍스트
131
이미지-추출
이미지 선택
130
증명서
칼라
단색
138
TEXT 작성
이미지-분해
132
RSI-적용
133
표 참조
137
문서화
Image(i)
134
키 적용
136
B(i)
image(i) 에 B(i)를 추출
135

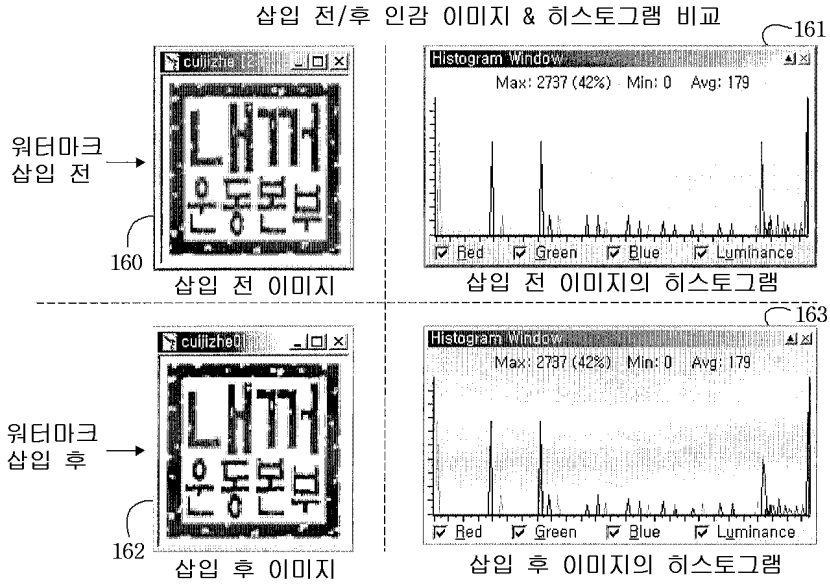
도면11


삽입 전/후 단색 이미지 & 히스토그램 비교
141
워터마크
삽입 전
140
삽입 전 이미지
Histogram Window
Max: 17054 (68%) Min: 0 Avg: 174
Red Green Blue Luminance
삽입 전 이미지의 히스토그램
143
워터마크
삽입 후
142
삽입 후 이미지
Histogram Window
Max: 17054 (68%) Min: 0 Avg: 174
Red Green Blue Luminance
삽입 후 이미지의 히스토그램

도면12


삽입 전/후 칼라 이미지 & 히스토그램 비교
151
워터마크
삽입 전
150
삽입 전 이미지
Histogram Window
Max: 1115 (1%) Min: 0 Avg: 128
Red Green Blue Luminance
삽입 전 이미지의 히스토그램
153
워터마크
삽입 후
152
삽입 후 이미지
Histogram Window
Max: 1264 (1%) Min: 0 Avg: 128
Red Green Blue Luminance
삽입 후 이미지의 히스토그램

도면13


삽입 전/후 인감 이미지 & 히스토그램 비교
161
워터마크
삽입 전
160
삽입 전 이미지
Histogram Window
Max: 2737 (42%) Min: 0 Avg: 179
Red Green Blue Luminance
삽입 전 이미지의 히스토그램
163
워터마크
삽입 후
162
삽입 후 이미지
Histogram Window
Max: 2737 (42%) Min: 0 Avg: 179
Red Green Blue Luminance
삽입 후 이미지의 히스토그램

도면14

단색 이미지 , 칼라 이미지, 인감 이미지들

도면15