(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2012-0097293
(43) 공개일자 2012년09월03일

(51) 국제특허분류(Int. Cl.)
G06F 17/30 (2006.01) *G06F 17/40* (2006.01)
(21) 출원번호 10-2011-0016795
(22) 출원일자 2011년02월24일
심사청구일자 없음

(71) 출원인
고려대학교 산학협력단
서울 성북구 안암동5가 1
(72) 발명자
이상진
서울특별시 노원구 중계동 363 대림벽산아파트 103동 1102호
전상준
서울특별시 성북구 인촌로27길 34-10, 인촌고시원 106호 (안암동5가)
(뒷면에 계속)
(74) 대리인
이건주, 김정훈

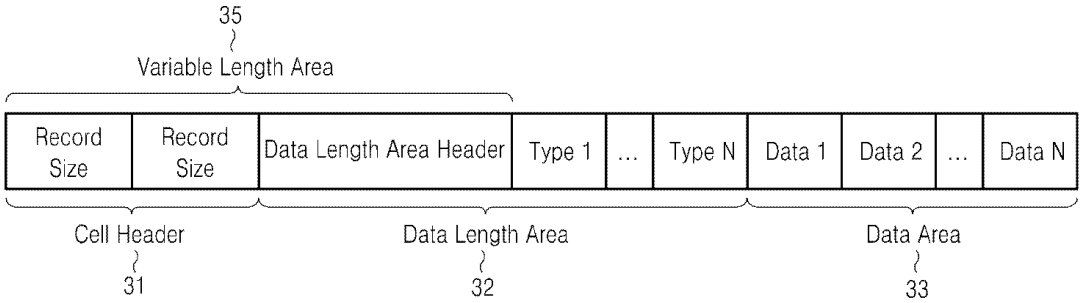
전체 청구항 수 : 총 18 항

(54) 발명의 명칭 **데이터베이스 복원 방법 및 장치**

(57) 요 약

본 발명에 따른 데이터베이스 복원 방법은 레코드된 데이터가 삭제된 영역을 비할당 영역으로 구분하여 관리하는 데이터베이스를 복원하는 방법에 있어서, 데이터의 레코드를 관리하는 정보를 포함한 데이터베이스 파일을 입력받는 과정과, 상기 데이터베이스 파일에 포함되며, 레코드된 데이터에 대한 구조를 지시하는 구성 정보를 확인하는 과정과, 상기 구성정보에 기초하여, 상기 레코드된 데이터가 삭제된 영역을 분석하는 과정과, 상기 레코드된 데이터가 삭제된 영역에 포함된 상기 레코드된 데이터를 검출하고 복원하는 과정을 포함한다.

대 표 도



(72) 발명자

변근덕

서울특별시 중랑구 겸재로 133, 204호 (면목동)

방제완

서울특별시 강서구 금낭화로 258, 방화10단지진로아파트 102동 207호 (방화동)

이 발명을 지원한 국가연구개발사업

과제고유번호	20100020634
부처명	한국연구재단
연구사업명	원천기술개발사업
연구과제명	디지털 정보 획득 기반기술 연구
주관기관	한국과학기술연구원
연구기간	2010.08.01 ~ 2011.07.31

특허청구의 범위

청구항 1

레코드된 데이터가 삭제된 영역을 비할당 영역으로 구분하여 관리하는 데이터베이스를 복원하는 방법에 있어서,

데이터의 레코드를 관리하는 정보를 포함한 데이터베이스 파일을 입력받는 과정과,

상기 데이터베이스 파일에 포함되며, 레코드된 데이터에 대한 구조를 지시하는 구성 정보를 확인하는 과정과,

상기 구성정보에 기초하여, 상기 레코드된 데이터가 삭제된 영역을 분석하는 과정과,

상기 레코드된 데이터가 삭제된 영역에 포함된 상기 레코드된 데이터를 검출하고 복원하는 과정을 포함함을 특징으로 하는 데이터베이스 복원 방법.

청구항 2

제1항에 있어서,

상기 구성 정보는 데이터베이스 파일의 종류를 지시하는 종류 정보와 테이블의 구조를 지시하는 구조 정보를 구비하는 스키마 테이블, 및 하위 페이지의 저장 위치를 지시하는 포인터 정보를 포함함을 특징으로 하는 데이터베이스 복원 방법.

청구항 3

제1항에 있어서, 상기 레코드된 데이터가 삭제된 영역을 분석하는 과정은,

상기 구조 정보로부터 비할당 영역에 포함된 셀의 길이를 확인하는 과정과,

상기 셀의 길이를 추정하는 과정과,

상기 구조 정보로부터 확인된 상기 셀의 길이와 상기 추정된 셀의 길이를 비교하여, 상기 셀에 상기 레코드된 데이터의 존재 여부를 확인하는 과정을 포함함을 특징으로 하는 데이터베이스 복원 방법.

청구항 4

제3항에 있어서, 상기 셀에 상기 레코드된 데이터의 존재 여부를 확인하는 과정은,

상기 구조 정보로부터 확인된 상기 셀의 길이와 상기 추정된 셀의 길이가 동일하면 상기 셀에 상기 레코드된 데이터가 존재하는 것으로 판단하는 과정을 특징으로 하는 데이터베이스 복원 방법.

청구항 5

제4항에 있어서,

상기 구조 정보로부터 확인된 상기 셀의 길이와 상기 추정된 셀의 길이가 동일하지 않음에 대응하여,

상기 구조 정보로부터 확인된 상기 셀을 분할하는 과정과,

상기 셀의 길이를 재 추정하는 과정과,

상기 분할된 셀의 길이와 상기 재 추정된 셀의 길이를 비교하여, 상기 분할된 셀에 상기 레코드된 데이터의 존재 여부를 확인하는 과정을 포함함을 특징으로 하는 데이터베이스 복원 방법.

청구항 6

제5항에 있어서,

상기 분할된 셀에 상기 레코드된 데이터의 존재 여부를 확인하는 과정은,

상기 분할된 셀의 길이와 상기 재 추정된 셀의 길이가 동일하면 상기 셀에 상기 레코드된 데이터가 존재하는 것으로 판단하는 과정을 포함함을 특징으로 하는 데이터베이스 복원 방법.

청구항 7

제6항에 있어서,

상기 분할된 셀에 상기 레코드된 데이터가 존재하는 것으로 판단한 후, 상기 분할된 셀의 다음에 위치하는 영역을 레코드된 데이터의 존재 여부를 확인하기 위한 분할된 셀로 재설정하는 과정을 더 포함함을 특징으로 하는 데이터베이스 복원 방법.

청구항 8

제6항에 있어서,

상기 분할된 셀의 길이와 상기 재 추정된 셀의 길이가 동일하지 않음에 대응하여, 상기 재 추정된 셀의 길이를 재설정하는 과정을 포함함을 특징으로 하는 데이터베이스 복원 방법.

청구항 9

제5항에 있어서, 상기 구조 정보로부터 확인된 상기 셀을 분할하는 과정은,

상기 셀의 길이를 순차적으로 증가시켜 분할하는 과정을 포함함을 특징으로 하는 데이터베이스 복원 방법.

청구항 10

레코드된 데이터가 삭제된 영역을 비할당 영역으로 구분하여 관리하는 데이터베이스를 복원하는 장치에 있어서,

데이터의 레코드를 관리하는 정보를 포함한 데이터베이스 파일을 입력받고, 상기 데이터베이스 파일에 포함되며, 레코드된 데이터에 대한 구조를 지시하는 구성 정보를 확인하는 스키마테이블 파서와,

상기 구성정보로부터 비할당 영역을 검출하는 비할당 영역 분석부와,

상기 비할당 영역에 구비된 적어도 하나의 셀의 길이를 추정하는 셀 길이 추정부와,

상기 비할당 영역 분석부와 셀 길이 추정부로부터 제공되는 정보를 이용하여 상기 레코드된 데이터가 삭제된 영역을 분석하여, 삭제된 데이터를 검출하는 데이터 검출부와,

상기 검출된 데이터를 복원하는 데이터 복원부를 포함함을 특징으로 하는 데이터베이스 복원 장치.

청구항 11

제10항에 있어서,

상기 구성 정보는 데이터베이스 파일의 종류를 지시하는 종류 정보와 테이블의 구조를 지시하는 구조 정보를 구비하는 스키마 테이블, 및 하위 페이지의 저장 위치를 지시하는 포인터 정보를 포함함을 특징으로 하는 데이터베이스 복원 장치.

청구항 12

제10항에 있어서, 상기 데이터 검출부는,

비할당 영역 분석부로부터 제공되는 상기 셀의 길이와 상기 셀 길이 추정부로부터 제공되는 상기 추정된 셀의 길이를 비교하여, 상기 셀에 상기 레코드된 데이터의 존재 여부를 확인하는 것을 특징으로 하는 데이터베이스 복원 장치.

청구항 13

제12항에 있어서, 상기 데이터 검출부는,

상기 구조 정보로부터 확인된 상기 셀의 길이와 상기 추정된 셀의 길이가 동일하면 상기 셀에 상기 레코드된 데이터가 존재하는 것으로 판단하는 것을 특징으로 하는 데이터베이스 복원 장치.

청구항 14

제13항에 있어서, 상기 데이터 검출부는,

상기 구조 정보로부터 확인된 상기 셀의 길이와 상기 추정된 셀의 길이가 동일하지 않음에 대응하여, 상기 구조 정보로부터 확인된 상기 셀을 분할하고, 상기 셀 길이 추정부로 상기 셀의 길이를 재추정할 것을 지시하고, 상기 분할된 셀의 길이와 상기 재 추정된 셀의 길이를 비교하여, 상기 분할된 셀에 상기 레코드된 데이터의 존재 여부를 확인하는 것을 포함함을 특징으로 하는 데이터베이스 복원 장치.

청구항 15

제14항에 있어서, 상기 데이터 검출부는,

상기 분할된 셀의 길이와 상기 재 추정된 셀의 길이가 동일하면 상기 셀에 상기 레코드된 데이터가 존재하는 것으로 판단하는 것을 특징으로 하는 데이터베이스 복원 장치.

청구항 16

제15항에 있어서,

상기 분할된 셀에 상기 레코드된 데이터가 존재하는 것으로 판단한 후, 상기 분할된 셀의 다음에 위치하는 영역을 레코드된 데이터의 존재 여부를 확인하기 위한 분할된 셀로 재설정하는 것을 특징으로 하는 데이터베이스 복원 장치.

청구항 17

제15항에 있어서,

상기 분할된 셀의 길이와 상기 재 추정된 셀의 길이가 동일하지 않음에 대응하여, 상기 재 추정된 셀의 길이를 재설정하는 것을 특징으로 하는 데이터베이스 복원 장치.

청구항 18

제14항에 있어서, 상기 데이터 검출부는,

상기 셀의 길이를 순차적으로 증가시켜 분할하는 것을 특징으로 하는 데이터베이스 복원 장치.

명 세 서

기 술 분 야

[0001] 본 발명은 데이터베이스 복원 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 대부분의 데이터가 디지털 형태로 변환되어 저장됨에 따라 방대한 양의 데이터를 효율적으로 관리하기 위해 데이터베이스의 사용이 꾸준히 증가하고 있다. 데이터베이스에 저장되는 정보는 시스템 운용에 필요한 핵심 정보인 경우가 많으며 포렌식 조사시 데이터베이스에서 의미 있는 정보를 찾게 될 가능성이 많아진다. 또한 데이터베이스에서 효율적인 저장 공간의 관리나 최신 데이터 유지를 위해 삭제 행위가 빈번하게 발생한다는 것을 고려할 때 데이터베이스에서 정상적으로 남아있는 데이터를 추출하는 것만큼 삭제된 데이터에 대한 정보를 획득하는 것도 중요하다.

[0003] SQLite는 작고 빠른 특징 때문에 간단하지만 많은 양의 정보를 체계적으로 저장할 필요가 있는 응용프로그램이나 임베디드 기기에 탑재된 소프트웨어에서 주로 사용되고 있다. 따라서 포렌식 조사에서 임베디드 기기의 증거를 수집할 경우 상당량의 데이터가 SQLite 데이터베이스 파일에 저장되어 있을 가능성이 많다. 따라서 SQLite 데이터베이스에서 삭제된 레코드를 복원하는 것은 상당히 중요하다.

[0004] SQLite를 사용하는 소프트웨어는 레코드를 삭제할 경우 각 소프트웨어별로 삭제된 데이터를 관리하는 정책이 다르다. 정책은 하기의 [표 1]과 같이 크게 세 가지로 분류된다.

표 1

[0005]

삭제된 데이터 관리 정책	복원 가능 여부
0으로 덮어씀	불가능

삭제된 영역 제거	불가능
비할당 영역으로 구분	가능

[0006] 첫째로 데이터를 삭제하면 0으로 덮어쓰는 경우가 있다. 이 경우 삭제된 영역을 파악하여도 데이터를 복구할 수 없다. 둘째로 삭제된 영역 자체를 제거하는 경우에는, 삭제행위가 발생했다는 흔적도 찾을 수 없다. 마지막으로 데이터 영역을 비할당 영역으로 구분하고 데이터는 그대로 남기는 경우가 있다. 이 경우, 데이터가 삭제된 영역을 단순하게 비할당 영역으로만 설정하게 되므로, 삭제된 후에도 파일에 남아있을 수 있고, 덮어 쓰이지 않은 경우 복구가 가능하다.

발명의 내용

해결하려는 과제

[0007] 본 발명은 전술한 점을 고려하여 안출된 것으로서, 데이터가 삭제된 영역을 비할당 영역으로 구분하여 관리하는 데이터베이스에서 삭제된 데이터를 효과적으로 복원할 수 있는 방법 및 장치를 제공하는데 그 목적이 있다.

과제의 해결 수단

[0008] 상기한 목적을 달성하기 위하여 본 발명의 일 측면에 따른 데이터베이스 복원 방법은 레코드된 데이터가 삭제된 영역을 비할당 영역으로 구분하여 관리하는 데이터베이스를 복원하는 방법에 있어서, 데이터의 레코드를 관리하는 정보를 포함한 데이터베이스 파일을 입력받는 과정과, 상기 데이터베이스 파일에 포함되며, 레코드된 데이터에 대한 구조를 지시하는 구성 정보를 확인하는 과정과, 상기 구성정보에 기초하여, 상기 레코드된 데이터가 삭제된 영역을 분석하는 과정과, 상기 레코드된 데이터가 삭제된 영역에 포함된 상기 레코드된 데이터를 검출하고 복원하는 과정을 포함한다.

[0009] 본 발명의 다른 측면에 따른 데이터베이스 복원 장치는, 데이터의 레코드를 관리하는 정보를 포함한 데이터베이스 파일을 입력받고, 상기 데이터베이스 파일에 포함되며, 레코드된 데이터에 대한 구조를 지시하는 구성 정보를 확인하는 스키마테이블 파서와, 상기 구성정보로부터 비할당 영역을 검출하는 비할당 영역 분석부와, 상기 비할당 영역에 구비된 적어도 하나의 셀의 길이를 추정하는 셀 길이 추정부와, 상기 비할당 영역 분석부와 셀 길이 추정부로부터 제공되는 정보를 이용하여 상기 레코드된 데이터가 삭제된 영역을 분석하여, 삭제된 데이터를 검출하는 데이터 검출부와, 상기 검출된 데이터를 복원하는 데이터 복원부를 포함한다.

발명의 효과

[0010] 본 발명에서 제시하는 데이터베이스 복원 방법 및 장치에 따르면, 데이터가 삭제된 영역을 비할당 영역으로 구분하여 관리하는 데이터베이스에서 삭제된 데이터를 효과적으로 복원할 수 있다.

도면의 간단한 설명

[0011] 도 1은 일반적인 데이터베이스 파일의 구조를 도시하는 개념도,

도 2는 일반적인 데이터베이스 파일에 구비된 페이지의 구조를 도시하는 개념도,

도 3은 도 2의 페이지에 구비된 셀의 구조를 도시하는 개념도,

도 4는 본 발명의 일 실시예에 따른 데이터베이스 복원 방법의 순서를 도시하는 흐름도,

도 5는 도 4의 45단계의 상세 순서를 예시하는 흐름도,

도 6은 도 5의 458단계의 상세 순서를 도시하는 흐름도,

도 7은 본 발명의 일 실시예에 따른 데이터베이스 복원 방법이 적용되는 전자장치의 구성을 예시하는 블록도,

도 8은 도 7의 전자장치에 구비된 제어부의 상세 구성을 도시하는 블록도.

발명을 실시하기 위한 구체적인 내용

[0012] 이하 본 발명에 따른 바람직한 실시예를 첨부한 도면을 참조하여 상세히 설명한다. 하기 설명에서는 구체적인 구성 소자 등과 같은 특정 사항들이 나타나고 있는데 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공

된 것일 뿐 이러한 특정 사항들이 본 발명의 범위 내에서 소정의 변형이나 혹은 변경이 이루어질 수 있음은 이 기술분야에서 통상의 지식을 가진 자에게는 자명하다 할 것이다.

[0013] 도 1은 일반적인 데이터베이스 파일의 구조를 도시하는 개념도이다. 도 1을 참조하면, 데이터베이스 파일은 데이터베이스 해더(11), 스키마 테이블(12), 및 적어도 하나의 페이지(13-1, …, 13-N)를 구비한다. 상기 데이터베이스 해더(11)는 데이터베이스 파일의 시그니처를 포함한 메타정보를 구비한다.

[0014] 상기 스키마 테이블(12)은 파일 내에 포함된 테이블의 구성에 대한 정보를 구비한다. 예컨대, 상기 스키마 테이블(12)은 스키마의 종류, 스키마의 이름, 루트 노드로서 마련되는 페이지의 포인터 정보, 및 스키마가 생성될 당시 전달되는 쿼리(Query)를 구비할 수 있다. 상기 스키마의 종류는 테이블, 인덱스, 및 트리거 중 어느 하나 일 수 있다. 또한, 예컨대, 상기 쿼리(Query)는 SQL 쿼리(Query) 일 수 있으며, 상기 SQL 쿼리(Query)는 스키마가 생성될 당시 SQLite로 전달되고, 필드의 타입, 필드 이름, 및 필드의 갯수 등의 정보를 포함할 수 있다.

[0015] 상기 페이지(13-1, …, 13-N)는 데이터베이스에 기록되는 데이터, 즉 레코드를 수록한다. 또한, 상기 페이지(13-1, …, 13-N)는 트리 형태의 구조로 이루어질 수 있다. 예컨대, 상기 페이지(13-1, …, 13-N)는 B-트리 형태의 구조로 이루어지며, 트리의 중간 단계에 위치하여 하위 페이지의 포인터에 대한 정보를 구비하는 인터널(internal) 페이지와, 하위 페이지의 포인터에 대한 정보를 구비하지 않고 상기 레코드를 구비하고 있는 립(Leaf) 페이지를 포함할 수 있다.

[0016] 도 2는 일반적인 데이터베이스 파일에 구비된 페이지의 구조를 도시하는 개념도이다. 도 2를 참조하면, 상기 페이지는 페이지 해더(21), 셀 오프셋 영역(22-1, 22-2, 22-3), 및 데이터 영역(23-1,23-2, 23-3, 24)을 포함한다.

[0017] 상기 페이지 해더(21)는 상기 페이지가 인터널(internal) 페이지 또는 립(Leaf) 페이지인지를 지시하는 페이지 플레그, 첫번째 비할당 영역의 오프셋, 첫번째 셀의 오프셋, 및 페이지 내에 포함된 셀의 갯수에 대한 정보를 포함한다.

[0018] 상기 데이터 영역(23-1,23-2, 23-3, 24)은 셀이 위치한 영역(23-1,23-2, 23-3)과 아직 셀이 할당되지 않아 초기 상태를 유지하는 비할당 공간 영역(24)을 포함한다.

[0019] 도 3은 도 2의 페이지에 구비된 셀의 구조를 도시하는 개념도이다. 도 3을 참조하면, 셀은 셀 해더(Cell Header)(31), 데이터 길이 영역(Data Length Area)(32), 및 데이터 영역(Data Area)(33)을 포함한다. 셀 해더(Cell Header)(31)는 레코드의 길이를 지시하는 레코드의 크기(Record Size)와, 기본키(Primary Key)를 구비한다. 그리고, 데이터 길이 영역(Data Length Area)(32)은 영역의 길이를 나타내는 데이터 길이 영역 해더(Data Length Area Header)와, 데이터 영역(33)에 위치하는 각 필드(Data 1, Data 2, ...Data N)에 대한 바이트 길이정보(Type 1, ... Type N)를 포함한다. 상기 레코드의 크기(Record Size), 기본키(Primary Key), 및 길이 영역 해더(Data Length Area Header)의 값은 가변 길이로 표현될 수 있음에 따라, 상기 레코드의 크기(Record Size), 기본키(Primary Key), 및 길이 영역 해더(Data Length Area Header)은 가변 길이 영역(Variable Length Area)(35)으로 설정한다.

[0020] 도 4는 본 발명의 일 실시예에 따른 데이터베이스 복원 방법의 순서를 도시하는 흐름도이다. 도 4를 참조하면, 우선 41단계에서, 데이터를 기록하고 있는 데이터베이스 파일을 입력받는다.

[0021] 다음으로, 43단계에서는, 상기 데이터베이스 파일의 구조를 참조하여, 데이터베이스 파일에 포함된 구성정보를 확인한다.

[0022] 구체적으로, 도 1에 도시된 바와 같이 구성된 데이터베이스 파일로부터 데이터베이스 해더(11), 스키마 테이블(12), 및 적어도 하나의 페이지(13-1, …, 13-N)를 확인한다. 그리고, 상기 스키마 테이블(12)을 확인하여, 스키마의 종류, 스키마의 이름, 루트 노드로서 마련되는 페이지의 포인터 정보, 및 스키마가 생성될 당시 전달되는 쿼리(Query)를 확인하고, 상기 적어도 하나의 페이지(13-1, …, 13-N)의 트리 형태의 구조를 확인한다. 그리고, 상기 적어도 하나의 페이지(13-1, …, 13-N)의 트리 형태의 구조에 기초하여, 립(Leaf) 페이지를 확인하고, 상기 립 페이지에 포함된 정보 즉, 페이지 해더(21), 셀 오프셋 영역(22-1, 22-2, 22-3), 및 데이터 영역(23-1,23-2, 23-3, 24)을 확인한다.

[0023] 다음으로, 45단계에서, 페이지 해더(21)와 셀 오프셋 영역(22-1, 22-2, 22-3)에 수록된 정보를 확인하여, 페이지 내에 설정된 비할당 영역을 검출하고, 상기 비할당 영역 내에 포함되며 데이터가 삭제된 영역인 셀 영역(23-1,23-2, 23-3)을 확인한다. 그리고, 상기 셀 영역 각각에 포함된 데이터를 분석하여, 상기 셀 영역(23-

1,23-2, 23-3, 24) 내의 데이터 영역(33)에 마련되는 각 필드(Data 1, Data 2, ...Data N)의 위치를 확인한다.

[0024] 마지막으로, 47단계에서, 데이터 영역(33)에 마련되는 각 필드(Data 1, Data 2, ...Data N)에 수록된 데이터를 추출하고, 추출된 데이터에 대한 복원을 진행한다. 즉 데이터 길이 영역(Data Length Area)(32)에 포함된 정보 중, 데이터 영역(33)에 위치하는 각 필드(Data 1, Data 2, ...Data N)에 대한 바이트 길이정보(Type 1, ... Type N)를 확인하여, 해당 필드의 바이트 길이 정보(Type 1, ... Type N)에 맞게 정수, 실수, 문자열, 바이너리 형태 등으로 필드의 데이터를 변환하여 출력한다.

[0025] 한편, 가변 길이 영역(Variable Length Area)(35)이 가변적인 길이로 설정됨에 따라, 전술한 45단계에서 레코드가 존재하는 데이터 영역(33)을 정확하게 파악하기 위해서는 가변 길이 영역(Variable Length Area)(35)에 포함된 각 영역(레코드의 크기, 기본키, 및 길이 영역 해더)의 길이를 정확하게 파악한 후, 데이터 영역(33)에 마련되는 각 필드(Data 1, Data 2, ...Data N)의 위치를 확인해야 한다.

[0026] 도 5는 도 4의 45단계의 상세 순서를 예시하는 흐름도이다.

[0027] 이하, 도 5를 참조하여, 45단계에서, 데이터 영역(33)에 마련되는 각 필드(Data 1, Data 2, ...Data N)가 위치한 영역을 검출하는 과정을 상세히 설명한다.

[0028] 우선, 451단계에서, 페이지 내에 설정된 비할당 영역에 포함된 각각의 셀의 길이를 입력받는다. 상기 각각의 셀의 길이는 상기 셀 오프셋 영역(22-1, 22-2, 22-3)에 수록된 셀의 오프셋 정보로부터 검출되어 입력될 수 있다.

[0029] 452단계에서는 상기 셀 영역(23-1,23-2, 23-3) 각각에 포함된 가변 길이 영역(35)을 최소 길이로 설정한다. 구체적으로, 가변 길이 영역(35)은 가변 길이 정수에 의해 추정될 수 있는데, 상기 가변 길이 정수는 최소 길이(m) 1부터 최대 길이(M)는 9까지의 값을 가질 수 있다.

[0030] 그리고, 가변 길이 영역(35)은 레코드의 크기(Record Size), 기본키(Primary Key), 및 길이 영역 해더(Data Length Area Header)를 구비한 3개의 영역으로 이루어지므로, 가변 길이 영역(35)의 전체 크기는 전술한 3개의 영역에 각각 상기 가변 길이 정수를 적용한 값으로 설정될 수 있다. 즉, 가변 길이 영역(35)의 크기는 하기의 수학식 1과 같이 연산될 수 있다.

수학식 1

[0031] $$VLA = 3 * N$$

[0032] VLA는 가변 길이 영역의 크기를 지시하고, N은 상기 가변 길이 정수 값을 지시한다. 결국, 상기 가변 길이 정수가 최소 길이(m) 1부터 최대 길이(M)는 9까지의 값을 가지므로, 가변 길이 영역의 크기는 3~27까지의 값을 가질 수 있다.

[0033] 다음으로, 453단계에서, 각각의 셀 영역의 전체 길이를 추정한다. 셀의 구조를 참조하여 셀 영역의 전체 길이는 하기의 수학식 2에 의해 연산될 수 있다.

수학식 2

[0034] $$f(x) = x + F + a + D$$

[0035] 수학식 1에서, f(x)는 셀의 전체 길이를 지시하고, x는 가변 길이 영역의 길이를 지시하고, F는 필드의 갯수를 지시하고, a는 TEXT와 BLOB가 56바이트 이상인 경우 발생하는 값을 지시하고, D는 레코드 데이터 영역의 길이를 지시한다.

[0036] 한편, 하나의 셀 영역 내에 하나의 레코드가 독립적으로 수록되어 있는 경우에는 상기 451단계에서 입력된 상기 셀의 길이와 상기 453단계에서 추정된 셀의 길이가 서로 동일한 값을 갖게 된다. 하지만, 하나의 셀 영역 내에 복수의 레코드가 존재하거나, 셀 영역이 비할당 영역으로 구분된 이후에 다른 데이터가 상기 비할당 영

역으로 구분된 셀 영역에 기록되어 데이터의 일부만 남아 있을 경우, 상기 451단계에서 입력된 상기 셀의 길이와 상기 453단계에서 추정된 셀의 길이가 서로 동일한 값을 갖지 않을 수 있다.

[0037] 하나의 셀 영역 내에 하나의 레코드가 독립적으로 수록되어 있는 경우 미리 정해진 셀의 최대 길이 이내에 비할당 영역이 존재하게 되지만, 하나의 셀 영역 내에 복수의 레코드가 존재하거나, 셀 영역이 비할당 영역으로 구분된 이후에 다른 데이터가 상기 비할당 영역으로 구분된 셀 영역에 기록되어 데이터의 일부만 남아있는 경우 미리 정해진 셀의 최대 길이 이내에 비할당 영역이 존재하지 않을 수 있다.

[0038] 따라서, 454단계에서는 상기 451단계에서 입력된 상기 셀의 길이와 상기 453단계에서 추정된 셀의 길이가 서로 동일한 값을 갖는지를 확인한다. 상기 입력된 셀의 길이와 상기 추정된 셀의 길이가 서로 동일한 값을 갖으면 455단계를 진행하고, 상기 입력된 셀의 길이와 상기 453단계에서 추정된 셀의 길이가 서로 동일하지 않으면, 456단계를 진행한다.

[0039] 455단계에서는 도 3에 도시된 셀의 구조에 기초하여, 상기 추정된 셀의 가변 길이 영역과 셀의 전체 길이를 확인하고, 상기 셀의 데이터 영역을 검출한다.

[0040] 한편, 456단계에서는, 상기 입력된 셀의 길이와 상기 추정된 셀의 길이가 서로 동일하지 않음에 따라, 상기 추정된 셀의 가변 길이 영역이 미리 정해진 최대 길이를 초과하는지를 확인한다. 상기 추정된 셀의 가변 길이 영역이 미리 정해진 최대 길이를 초과하면 458단계를 진행하고, 상기 추정된 셀의 가변 길이 영역이 미리 정해진 최대 길이를 초과하지 않으면 457단계를 진행한다.

[0041] 457단계에서는 상기 셀의 가변 길이를 재설정한다. 즉, 상기 가변 길이 정수를 최소 길이(m)부터 순차적으로 1만큼 증가시켜 상기 셀의 가변 길이를 재설정한다. 그리고, 457단계에서 상기 셀의 가변 길이가 재설정되면 다시 453단계를 진행한다. 이에 따라, 453, 454, 456, 및 457단계는 상기 가변 길이 정수가 최대 길이(M) 값으로 설정될 때까지 반복적으로 수행되며, 상기 가변 길이 정수가 최대 길이(M) 값을 초과하여 설정되면 추정된 셀의 가변길이가 미리 정해진 최대 길이를 초과하게 되므로, 456단계에서 458단계로 진행되게 된다.

[0042] 한편, 상기 추정된 셀의 가변 길이 영역이 미리 정해진 최대 길이를 초과하면, 하나의 셀 영역 내에 복수의 레코드가 존재하거나, 셀 영역이 비할당 영역으로 구분된 이후에 다른 데이터가 상기 비할당 영역으로 구분된 셀 영역에 기록되어 데이터의 일부만 남아있을 수 있으므로, 458단계에서는, 하나의 셀 영역 내에 복수의 레코드가 존재하는지를 확인한다.

[0043] 도 6은 도 5의 458단계의 상세 순서를 도시하는 흐름도이다. 이하, 도 6을 참조하여, 상기 458단계의 구체적인 동작을 설명한다.

[0044] 우선, 4580단계에서는, 비할당 영역의 길이를 설정한다. 비할당 영역의 길이의 초기 값은 상기 451단계에서 확인된 제1 비할당 영역에 포함된 필드의 갯수를 고려하여 설정할 수 있다.

[0045] 4581 단계에서는 상기 비할당 영역의 길이로 상기 451단계에서 확인된 비할당 영역을 분할한 후, 분할된 비할당 영역 중, 가장 앞에 있는 영역을 제2 비할당 영역으로 설정한다.

[0046] 다음으로, 4582단계에서는, 452단계에서와 같이, 상기 셀 영역(23-1,23-2, 23-3) 각각에 포함된 가변 길이 영역(35)을 최소 길이로 설정한다. 예컨대, 가변 길이 영역(35)을 최소값인 3으로 설정한다. 그리고, 4583단계에서는, 전술한 453단계에서와 같이, 상기의 수학식 2에 의해 셀의 전체 길이를 추정한다.

[0047] 4584단계에서는, 상기 4581단계에서 설정된 상기 제2 비할당 영역의 길이와 상기 5483단계에서 추정된 셀의 전체 길이가 서로 동일한 값을 갖는지를 확인한다. 상기 제2 비할당 영역의 길이와 상기 추정된 셀의 길이가 서로 동일한 값을 갖으면 4586단계를 진행하고, 상기 제2 비할당 영역의 길이와 상기 5483단계에서 추정된 셀의 길이가 서로 동일하지 않으면, 4585단계를 진행한다.

[0048] 4585단계에서는 전술한 457단계에서와 같이 상기 셀의 가변 길이를 재설정한다. 즉, 상기 가변 길이 정수를 최소 길이(m)부터 순차적으로 1만큼 증가시켜 상기 셀의 가변 길이를 재설정한다. 그리고, 4585단계에서 상기 셀의 가변 길이가 재설정되면 다시 4583단계를 진행한다. 이에 따라, 4583, 4584, 및 4585단계는 상기 4583단계에서 추정되는 셀의 전체 길이가 상기 4581단계에서 설정된 제2 비할당 영역의 길이와 동일해질 때까지 반복적으로 수행된다.

[0049] 한편, 4586단계에서는, 상기 4581단계에서 설정된 제2 비할당 영역을 데이터 영역으로 확인한다.

[0050] 그리고, 4587단계에서는, 상기 451단계에서 확인된 제1 비할당 영역 내에 상기 데이터 영역으로 확인된 영역

이외에 잔여 영역이 존재하는 지를 확인한다. 만약 잔여 영역이 존재하는 경우 4588단계를 진행한다. 4588단계에서는 상기 잔여 영역 중, 상기 데이터 영역의 다음에 위치하며 상기 4580단계에서 설정된 제2 비할당 영역의 길이 만큼에 대응하는 영역을 제2 비할당 영역으로 재설정한다.

[0051] 4582, 4583, 4584, 4585, 4586, 4587, 및 4588단계는 상기 451단계에서 확인된 제1 비할당 영역 내에 상기 데이터 영역으로 확인된 영역 이외에 잔여 영역이 존재하지 않을 때까지 반복적으로 수행하여, 상기 451단계에서 확인된 제2 비할당 영역 전체에 대해 데이터 영역을 확인한다.

[0052] 상기 제1 비할당 영역에 복수의 셀 영역이 존재하는 경우, 상기 셀 영역의 길이는 다양한 값을 가질 수 있으므로, 상기 제1 비할당 영역의 범위 내에서 제2 비할당 영역의 길이를 순차적으로 증가시켜 설정하는 것이 바람직하다. 따라서, 4589단계에서는 상기 제1 비할당 영역의 길이와 제2 비할당 영역의 길이를 서로 비교하여 상기 제2 비할당 영역의 길이가 상기 제1 비할당 영역의 길이를 초과하는 지를 확인한다.

[0053] 그리고, 상기 제2 비할당 영역의 길이가 상기 제1 비할당 영역의 길이를 초과하지 않을 경우 4590단계를 진행하여, 상기 제2 비할당 영역의 길이를 순차적으로 증가시킨다. 따라서, 4581, 4582, 4583, 4584, 4585, 4586, 4587, 4588, 4589, 및 4590 단계는 상기 제2 비할당 영역의 길이가 상기 제1 비할당 영역의 길이에 도달할 때까지 반복적으로 수행한다.

[0054] 도 7은 본 발명의 일 실시예에 따른 데이터베이스 복원 방법이 적용되는 전자장치의 구성을 예시하는 블록도이다.

[0055] 상기 데이터베이스 복원 방법이 적용되는 전자장치는 키 입력부(71), 디스플레이부(72), 메모리부(73), 외부 입/출력 인터페이스(74), 및 제어부(77)를 포함한다.

[0056] 키 입력부(71)는 사용자로부터 데이터를 입력받기 위한 장치로써, 숫자 및 문자 정보를 입력하기 위한 키들 및 각종 기능을 설정하기 위한 기능키들을 포함하며, 상기 키들은 키패드로 구현되거나, 터치 스크린과 오버랩된 디스플레이를 통해 입력받을 키들을 표시하고, 터치가 입력된 영역에 대응하는 키를 입력받는 터치 스크린 기반의 키 입력 인터페이스로 구현될 수 있다.

[0057] 디스플레이(72)는 액정표시장치(Liquid Crystal Display: LCD) 등과 같은 표시장치로 이루어질 수 있으며, 제어부(77)의 제어 하에 해당 전자장치의 각종 동작 상태에 대한 메시지를 비롯하여, 어플리케이션의 구동에 따라 발생되는 데이터 등을 표시한다.

[0058] 메모리부(73)는 어플리케이션의 구동에 필요한 데이터, 예컨대 데이터베이스 복구를 위한 어플리케이션의 실행을 위한 데이터 등을 저장한다. 특히, 메모리부(73)는 데이터베이스 복구를 위한 어플리케이션의 실행을 통해 복구되는 데이터를 저장할 수 있다.

[0059] 외부 입/출력 인터페이스(74)는 복구 데이터베이스(75)를 연결할 수 있도록 마련되고, 상기 복구 데이터베이스(75)에 수록된 데이터를 제어부(77)에 제공할 수 있도록 구비된다.

[0060] 제어부(77)는 상기 각 기능부들의 동작을 총괄적으로 전자장치의 전반적인 동작을 제어하는 기능을 수행한다. 즉, 제어부(77)는 키 입력부(71)를 통해 입력되는 신호에 따른 처리를 수행하고, 어플리케이션의 실행에 의해 발생되는 데이터를 메모리(73)에 저장하거나, 디스플레이부(72)를 통해 출력하도록 제어한다.

[0061] 특히, 제어부(77)는 데이터베이스 복구를 위한 어플리케이션의 실행을 통해, 외부 입/출력 인터페이스(74)에 연결된 복구 데이터베이스(75)에 수록된 데이터를 검출하고, 검출된 데이터를 복구하는 기능을 수행한다.

[0062] 도 8은 도 7의 전자장치에 구비된 제어부의 상세 구성을 도시하는 블록도이다. 도 8을 참조하면, 제어부(77)는 스키마테이블 파서(81), 비할당 영역 길이 확인부(82), 셀 길이 추정부(83), 데이터 검출부(84), 및 데이터 복원부(85)를 포함한다.

[0063] 스키마테이블 파서(81)는, 복구 데이터베이스(75)로부터 데이터베이스 파일에 포함된 구성정보를 확인한다. 구체적으로, 도 1에 도시된 바와 같이 구성된 데이터베이스 파일로부터 데이터베이스 해더(11), 스키마 테이블(12), 및 적어도 하나의 페이지(13-1, …, 13-N)를 검출한다. 그리고, 상기 스키마 테이블(12)을 확인하여, 스키마의 종류, 스키마의 이름, 루트 노드로서 마련되는 페이지의 포인터 정보, 및 스키마가 생성될 당시 전달되는 쿼리(Query)를 확인하고, 상기 적어도 하나의 페이지(13-1, …, 13-N)의 트리 형태의 구조를 확인한다. 그리고, 상기 적어도 하나의 페이지(13-1, …, 13-N)의 트리 형태의 구조에 기초하여, 립(Leaf) 페이지를 확인하고, 상기 립 페이지에 포함된 정보 즉, 페이지 해더(21), 셀 오프셋 영역(22-1, 22-2, 22-3),

및 데이터 영역(23-1,23-2, 23-3, 24)을 확인한다.

[0064] 비할당 영역 길이 확인부(82)는 페이지 해더(21)와 셀 오프셋 영역(22-1, 22-2, 22-3)에 수록된 정보를 확인하여, 페이지 내에 설정된 비할당 영역을 검출하고, 상기 비할당 영역 내에 포함되며 데이터가 삭제된 영역인 셀 영역(23-1,23-2, 23-3)을 확인한다. 그리고, 상기 비할당 영역 내에 포함되며 데이터가 삭제된 영역인 셀 영역(23-1,23-2, 23-3)의 길이를 확인하여, 데이터 검출부(84)에 제공한다.

[0065] 셀 길이 추정부(83)는 셀의 가변 길이 영역을 추정하여, 셀의 전체 길이를 추정하는 기능을 수행한다. 예컨대, 상기 셀 길이 추정부(83)는 도 5의 452, 453 및 457 단계의 수행을 처리한다.

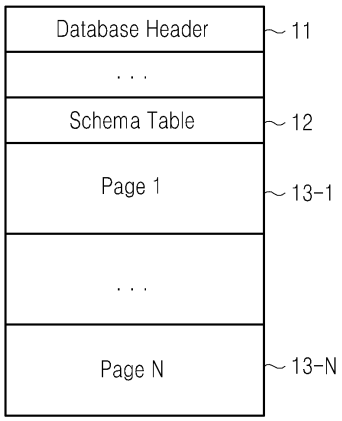
[0066] 데이터 검출부(84)는 비할당 영역 길이 확인부(82)와 셀 길이 추정부(83)로부터 제공되는 데이터를 사용하여, 상기 셀 영역 각각에 포함된 데이터를 분석하여, 상기 셀 영역(23-1,23-2, 23-3, 24) 내의 데이터 영역(33)에 마련되는 각 필드(Data 1, Data 2, ...Data N)의 위치를 확인한다. 그리고, 데이터 영역(33)에 마련되는 각 필드(Data 1, Data 2, ...Data N)에 수록된 데이터를 검출한다. 구체적으로, 데이터 검출부(84)는 비할당 영역 길이 확인부(82)로부터 제공되는 비할당 영역 길이 정보와 셀 길이 추정부(83)로부터 제공되는 추정된 셀 전체 길이 정보를 비교하는 454단계를 수행하고, 추정된 셀의 가변 길이 영역의 정보를 확인하는 456단계를 수행할 수 있다. 또한, 456단계의 수행 결과에 따라, 상기 셀 길이 추정부(83)로 셀의 전체 길이의 재설정을 요청하거나, 비할당 영역을 분할하여 데이터 영역을 확인하는 등의 동작을 수행한다.

[0067] 데이터 복원부(85)는 데이터 검출부로부터 검출된 데이터, 즉 데이터 영역(33)에 위치하는 각 필드(Data 1, Data 2, ...Data N)에 대한 바이트 길이정보(Type 1, ... Type N)를 확인하여, 해당 필드의 바이트 길이 정보(Type 1, ... Type N)에 맞게, 상기 각 필드(Data 1, Data 2, ...Data N)에 수록된 데이터를 정수, 실수, 문자열, 바이너리 형태 등으로 필드의 데이터를 변환하여 데이터를 복원한다.

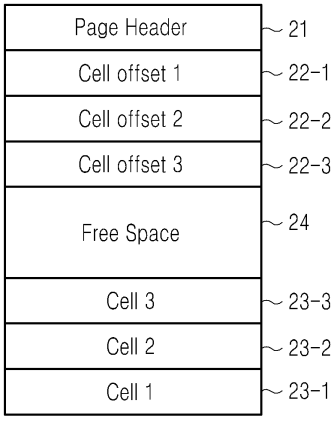
[0068] 이상에서 본 발명은 비록 한정된 실시예와 도면에 의해 설명되었으나, 본 발명은 이것에 의해 한정되지 않으며 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 다양한 수정 및 변형이 가능함은 물론이다.

도면

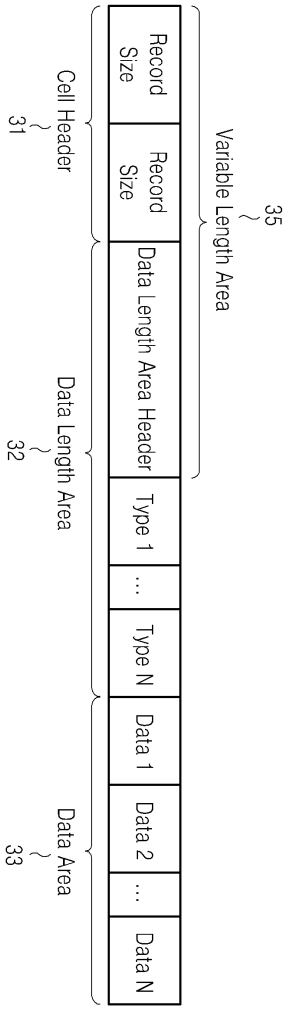
도면1



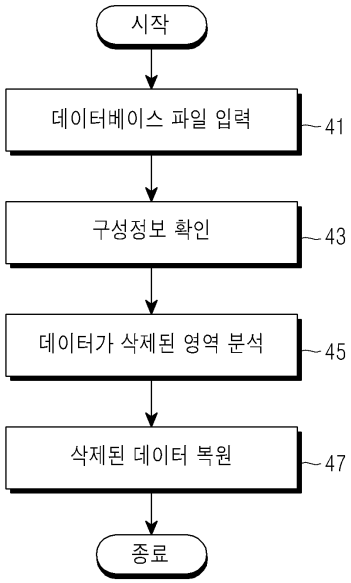
도면2


Page Header
21
Cell offset 1
22-1
Cell offset 2
22-2
Cell offset 3
22-3
Free Space
24
Cell 3
23-3
Cell 2
23-2
Cell 1
23-1

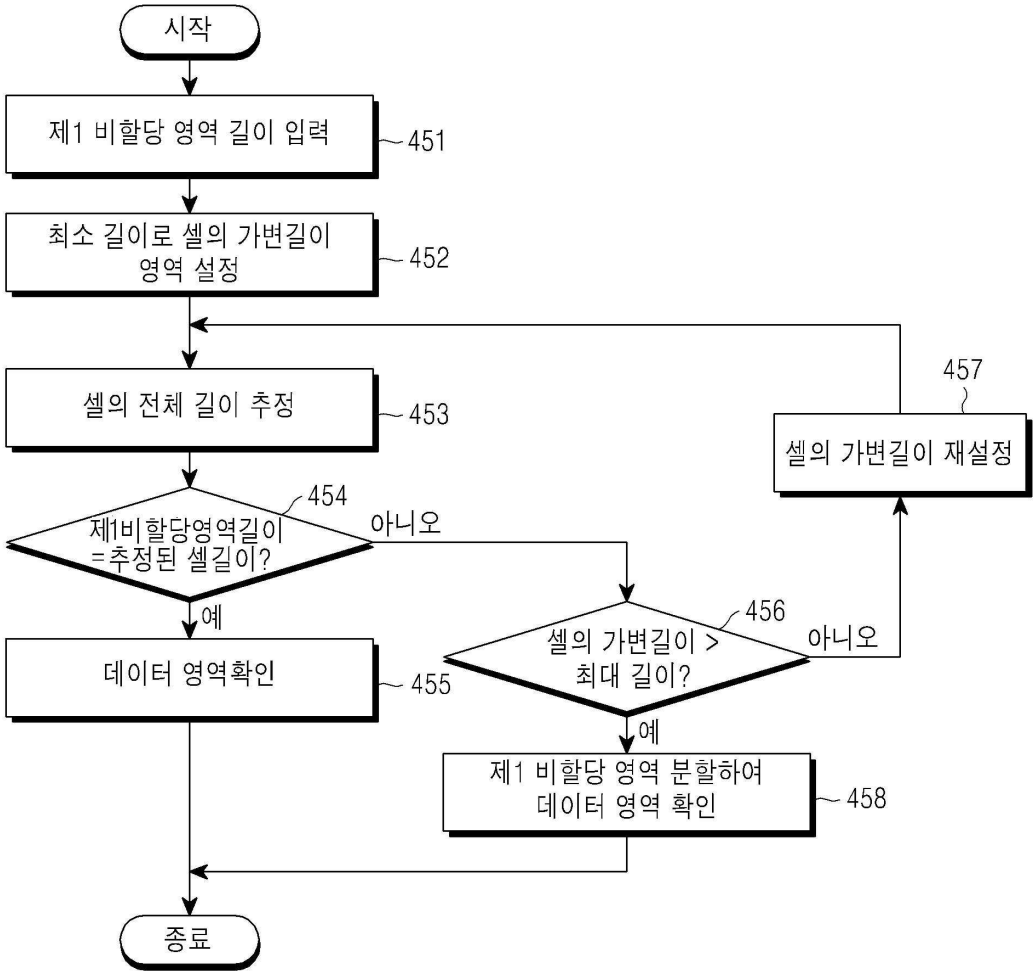
도면3


Record Size
Record Size
Data Length Area Header
Type 1
...
Type N
Data 1
Data 2
...
Data N
Variable Length Area
35
Cell Header
31
Data Length Area
32
Data Area
33

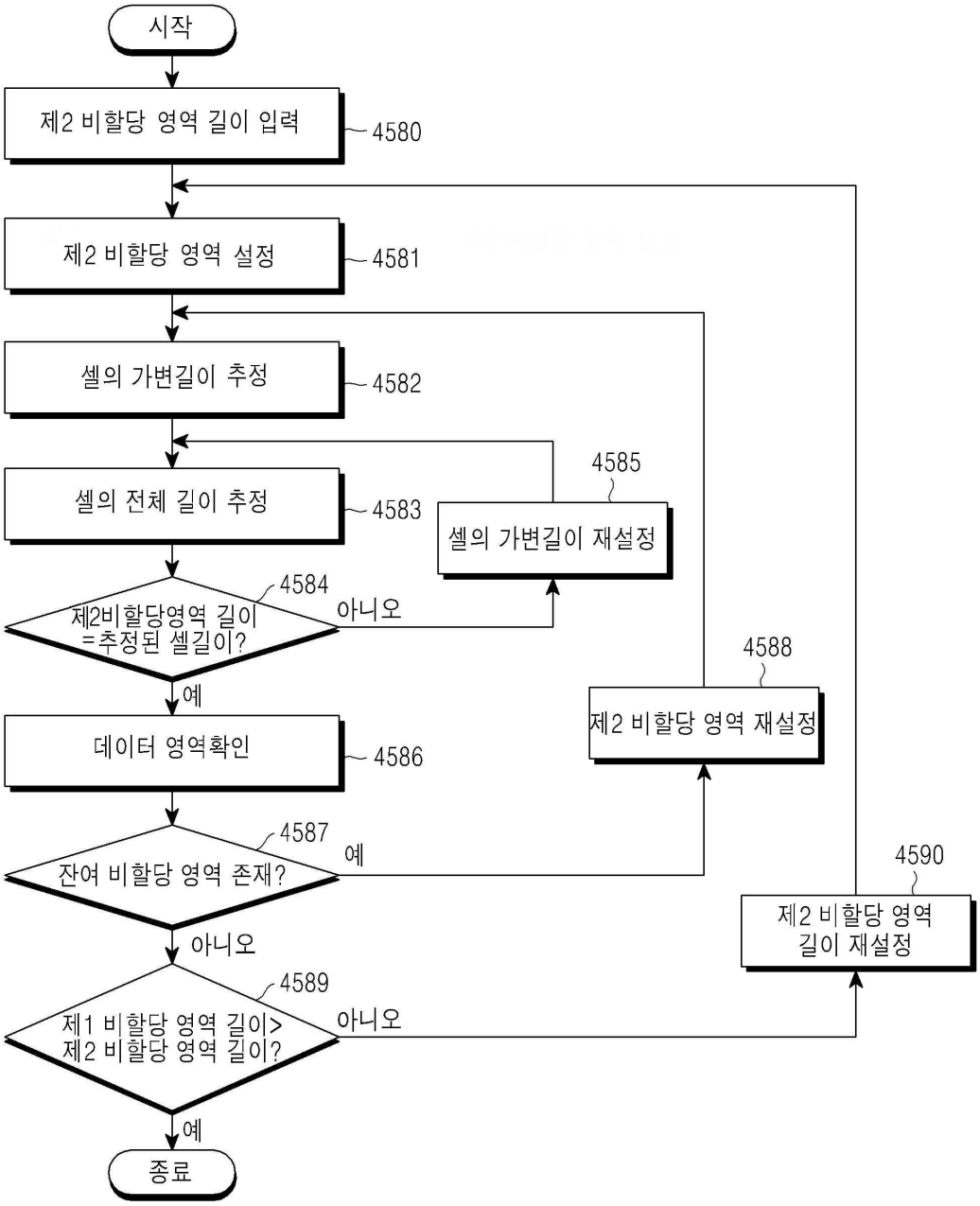
도면4


시작
데이터베이스 파일 입력
41
구성정보 확인
43
데이터가 삭제된 영역 분석
45
삭제된 데이터 복원
47
종료

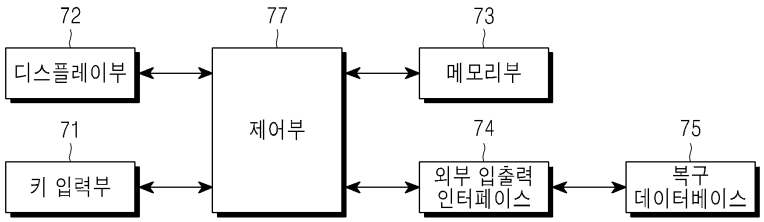
도면5


시작
제1 비할당 영역 길이 입력
451
최소 길이로 셀의 가변길이 영역 설정
452
셀의 전체 길이 추정
453
제1비할당영역길이 =추정된 셀길이?
454
아니오
예
데이터 영역확인
455
셀의 가변길이 > 최대 길이?
456
아니오
예
셀의 가변길이 재설정
457
제1 비할당 영역 분할하여 데이터 영역 확인
458
종료

도면6


시작
제2 비할당 영역 길이 입력
4580
제2 비할당 영역 설정
4581
셀의 가변길이 추정
4582
셀의 전체 길이 추정
4583
4585
셀의 가변길이 재설정
4584
제2비할당영역 길이
=추정된 셀길이?
아니오
예
데이터 영역확인
4586
4588
제2 비할당 영역 재설정
4587
잔여 비할당 영역 존재?
예
아니오
4590
제2 비할당 영역
길이 재설정
4589
제1 비할당 영역 길이>
제2 비할당 영역 길이?
아니오
예
종료

도면7


72
디스플레이부
77
제어부
73
메모리부
71
키 입력부
74
외부 입출력
인터페이스
75
복구
데이터베이스

도면8

