



광학 저장 매체 상에 LDPC 코드를 이용한 데이터의 분산 기록 방법

A Technique of Distributed Writing on Optical Disc with LDPC Codes

저자 (Authors)	김태웅, 류준길, 박찬익 Taewoong Kim, Junkil Ryu, Chanik Park
출처 (Source)	한국정보과학회 학술발표논문집 35(1A) , 2008.6, 327-328 (2 pages)
발행처 (Publisher)	한국정보과학회 KOREA INFORMATION SCIENCE SOCIETY
URL	http://www.dbpia.co.kr/Article/NODE01010444
APA Style	김태웅, 류준길, 박찬익 (2008). 광학 저장 매체 상에 LDPC 코드를 이용한 데이터의 분산 기록 방법. 한국정보과학회 학술발표논문집, 35(1A), 327-328.
이용정보 (Accessed)	포항공과대학교 141.223.121.100 2016/05/09 16:41 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다.

이 자료를 원저작자와의 협의 없이 무단게재 할 경우, 저작권법 및 관련법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

The copyright of all works provided by DBpia belongs to the original author(s). Nurimedia is not responsible for contents of each work. Nor does it guarantee the contents.

You might take civil and criminal liabilities according to copyright and other relevant laws if you publish the contents without consultation with the original author(s).

광학 저장 매체 상에 LDPC 코드를 이용한 데이터의 분산 기록 방법

김태웅[○] 류준길 박찬익

포항공과대학교

{ehoto[○], lancer, cipark}@postech.ac.kr

A Technique of Distributed Writing on Optical Disc with LDPC Codes

Taewoong Kim[○] Junkil Ryu Chanik Park

POSTECH

광학 저장 매체는 용량 대비 가격이 저렴하고 기술의 발달로 인해 집적도 또한 증가해 현재 블루레이 디스크의 경우 한 장이 25GB, 듀얼 레이어의 경우 50GB의 용량을 가지고 있다. 또한 테이프에 비해 온도와 습도와 같은 환경적인 요인으로 인한 영향이 적어 데이터 장기 보관에 유리하고 랜덤한 접근이 가능하다는 장점을 가지고 있다. 그러나 데이터 기록면이 외부에 노출되어있는 광학 저장 매체의 특성상 스크래치와 같은 표면의 손상으로 쉽게 데이터를 손실할 수 있다. 데이터 손실을 방지하기 위해 광학 저장 매체 내부에는 데이터를 기록할 시 여러 개의 섹터 단위로 오류 정정 부호(ECC: Error Correcting Codes) 알고리즘을 통해 기록한다.[1][2][3]. 이와 같은 광학 저장 매체의 자체 ECC 메커니즘을 통해 CD는 약 2.4 mm, DVD는 약 6 mm, Blu-ray Disc는 약 7 mm 길이 이내의 표면 손상이 발생한 경우에는 데이터 복원이 가능하나 그 이상의 표면 손상이 발생 시 데이터 손실이 발생할 수 있어 제한적인 신뢰도를 제공하고 있다. 본 논문에서는 저장할 데이터에 erasure 코드의 한 종류인 LDPC(Low Density Parity Check)[4] 코드를 적용하여 중복 데이터를 생성하고 중복 데이터들을 광학 저장 매체 상에 요구 신뢰도를 충족시켜줄 수 있는 간격을 갖도록 고루 분산시켜 배치하는 방법을 이용해 신뢰도를 향상시키는 방법을 제안한다. 전체적인 구조는 [그림 1]와 같이 두 개의 계층으로 구성되어 있다.

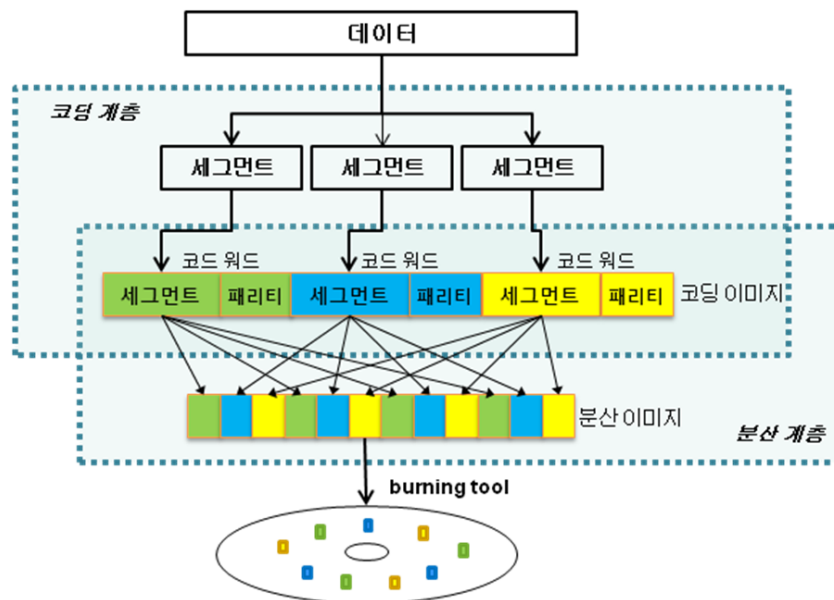


그림 1. 사용자 데이터의 처리 과정

첫 번째 계층인 코딩 계층에서는 데이터를 보호하기 위해 LDPC 코드를 사용한다. 사용자의 데이터는 일정한 개수의 블록으로 구성된 세그먼트 단위로 처리된다. 코딩 계층에서는 세그먼트 단위로 하나씩 읽어 들어 LDPC 코드를 통해 패리티 블록을 생성하고 세그먼트 뒤에 패리티 블록들을 붙여 하나의 코드 워드를 생성한다. 이렇게 생성된 코드 워드 내에 일부 데이터가 손실되더라도 LDPC의 디코딩 알고리즘을 통해 복구가 가능해진다. 사용자의 데이터는 코딩 계층을 통해 여러 개의 코드 워드로 만들어져 코딩 이미지로 하드 디스크에 임시로 저장된다.

고밀도로 집적된 광학 저장 매체에서는 작은 스크래치조차 수많은 연속적인 데이터 손실을 야기한다. 따라서 하나

의 스크래치만으로도 코드 워드의 상당 부분을 잃어서 디코딩이 불가능해 질 수 있다. 두 번째 계층인 분산 계층에서는 코드 워드를 광학 저장 매체에 연속적인 영역에 배치하지 않고 미디어 전체에 고루 분포 시켜 하나의 스크래치는 코드 워드의 일부분만을 손상 시킬 수 있도록 하였다. 분산 계층은 코딩 이미지에서 코드 워드를 하나씩 읽어 들여와 코드 워드의 각 블록들이 일정 간격을 두고 배치되도록 분산 이미지에 저장을 한다. 분산 이미지 파일 상에서는 동일한 블록 간격을 두고 코드 워드의 블록들이 배치되어있어도 데이터를 나선형으로 기록하는 광학 저장 매체의 특성에 따라 실제로 매체 상에 기록 시 코드 워드의 블록들은 각기 다른 트랙 간의 간격을 갖으며 위치하게 된다. 나선의 둘레가 외곽으로 갈수록 길어져 외곽으로 갈수록 코드 워드 블록들은 적은 트랙 간 간격을 두며 배치가 된다. 따라서 최외곽에 배치되는 2개 블록의 트랙 간격을 사용자가 요구하는 트랙 거리가 되도록 블록 간격을 설정하면 안쪽 트랙의 블록들도 사용자가 요구하는 트랙 거리 이상의 간격을 두고 배치가 된다. 데이터가 기록되는 최외곽의 트랙과 그 트랙에서 사용자가 요구하는 d의 거리만큼 떨어진 트랙 사이에 존재하는 블록의 개수를 블록 간격 g로 이용한다. 이를 계산하는 방법은 다음과 같다.

$$\text{블록 간격 } g = \left\lfloor \frac{\text{SpiralLen}(\text{OuterRadius})}{\text{BlockLength}} \right\rfloor - \left\lfloor \frac{\text{SpiralLen}(\text{OuterRadius} - d)}{\text{BlockLength}} \right\rfloor$$

OuterRadius는 최 외각 트랙이 위치하는 매체 상의 반지름이며 BlockLength는 광학 저장 매체에서 사용되는 블록의 물리적인 길이를 뜻한다. SpiralLen(r)은 광학 저장 매체의 중심으로부터 해당 반지름 r까지의 물리적인 나선 거리나 타내며 이는 Archimedes의 나선 공식을 사용하여 아래와 같이 계산할 수 있다.

```
float SpiralLen (r)
{
    radian =  $\frac{2 \pi r}{\text{TrackPitch}}$ 
    return  $\frac{1}{2} \frac{\text{TrackPitch}}{2 \pi} \left( \text{radian} \sqrt{1 + \text{radian}^2} + \text{Log} \left( \text{radian} + \sqrt{1 + \text{radian}^2} \right) \right)$ 
}
```

TrackPitch는 광학 저장 매체에서 인접 트랙 간의 물리적 거리를 뜻하며 radian은 반지름 r까지 도달하기 위해 필요한 나선상에서의 회전 각도를 뜻한다. 위와 같은 방법으로 계산된 블록 간격 g를 이용하여 코드 워드의 블록을 배치하게 되면 하나의 코드 워드에 속하는 블록들은 d 트랙 간격 이상을 가지며 배치된다. 따라서 광학 저장 매체에 d 두께의 손상이 발생해도 두 개 이상의 코드 워드 블록을 손상시키지 않게 된다.

평가를 위해 인접 트랙 간 간격이 0.74μm이고 32Kbyte의 크기의 ECC 블록을 가지며 약 4.3 Gbyte의 용량(143656개의 ECC 블록)을 갖는 120mm DVD-R을 기준으로 분석하였다. 12개의 블록의 크기를 갖는 코드 워드를 사용했으며 이중 6개를 패리티 블록으로 사용하였다. 이러한 설정의 LDPC 코드를 사용하여 코드 워드를 생성했을 시 코드 워드는 3개의 블록 손실까지만 데이터를 복구할 수 있다. 그러나 2mm 이상의 트랙 간 간격을 갖도록 코드 워드의 블록들을 분산 배치했을 시 블록들은 서로 11971개의 블록 간격으로 배치되고 35913개의 연속적인 블록 손상까지 데이터를 복구할 수 있었다. 35913개의 연속적인 블록 손상은 11971개의 코드 워드로 분산되어 하나의 코드 워드에는 3개의 블록만이 손상되기 때문이다. 이와 같이 LDPC 코드를 이용하여 코드 워드를 생성하고 코드 워드를 광학 저장 매체 전체에 고루 분산시켜 저장 시키는 기법을 이용하여 추가적인 신뢰도를 제공할 수 있었다. 또한 사용자가 지정하는 수치 이상의 트랙 간 간격을 유지하며 코드 워드 블록을 배치하도록 하여 사용자가 요구하는 정도의 신뢰도를 제공하였다.

[Acknowledgement]

본 연구는 2008년도 두뇌한국 21 사업과 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구 결과로 수행되었음. (IITA-2008-C1090-0801-0045)

[참고 문헌]

[1] ISO International Standard, "Information technology – Data interchange on read-only 120 mm optical data disks (CD-ROM)", Reference number: ISO/IEC 10149:1995(E).
 [2] European Computer Manufacturers Association (ECMA), "Standard ECMA-359: 80 mm (1,46 Gbytes per side) and 120 mm (4,70 Gbytes per side) DVD Recordable Disk (DVD-R)," ECMA International, 1st Edition, December 2004.
 [3] Blu-ray Disc Association, "Blu-ray Disc Recordable Format", Part 1: Physical specification, February 2006.
 [4] R. G. Gallager, Low Density Parity-Check Codes. MIT Press, Cambridge, MA, 1963.