(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0152414
(43) 공개일자 2023년11월03일

(51) 국제특허분류(Int. Cl.)
G06F 17/16 (2006.01) *G06F 5/01* (2006.01)
G06F 7/523 (2006.01)
(52) CPC특허분류
G06F 17/16 (2013.01)
G06F 5/012 (2013.01)
(21) 출원번호 10-2022-0052145
(22) 출원일자 2022년04월27일
심사청구일자 2022년04월27일

(71) 출원인
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
김지훈
서울특별시 서초구 서초대로38길 12, 102동 603호(서초동, 마제스타시티, 힐스테이트 서리풀)
김휘진
경기도 김포시 유현로 19, 115동 902호(풍무동, 유현마을신동아아파트)
(74) 대리인
리앤목특허법인

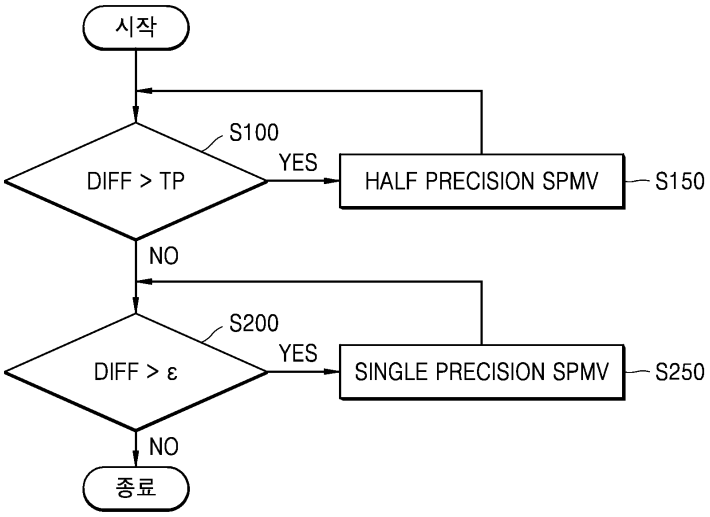
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 **페이지랭크 희소행렬 곱셈 연산을 위한 연산 장치 및 방법**

(57) 요 약

페이지랭크 희소행렬 곱셈 연산을 위한 연산 장치 및 방법을 위하여, 본 발명에 따른 연산 장치는, 페이지랭크 알고리즘의 희소행렬 연산을 수행하는 적어도 하나의 프로세서를 포함하는 연산 장치에 있어서, 상기 프로세서는, 상기 페이지랭크 알고리즘으로부터 페이지랭크 벡터를 획득하는 벡터 획득 모듈, 상기 페이지랭크 벡터의 유클리드 거리값을 연산하는 거리값 연산 모듈, 및 상기 유클리드 거리값을 기초로 상기 희소행렬 연산의 부동소수점 정밀도를 변환하는 정밀도 변환 모듈을 포함하는 연산 장치를 제공한다.

대 표 도 - 도1



(52) CPC특허분류
G06F 7/523 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711152413
과제번호	2021-0-00853-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	신개념PIM반도체선도기술개발(R&D)
연구과제명	PIM 활용을 위한 SW 플랫폼 개발
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2022.01.01 ~ 2022.12.31

명 세 서

청구범위

청구항 1

페이지랭크 알고리즘의 희소행렬 연산을 수행하는 적어도 하나의 프로세서를 포함하는 연산 장치에 있어서,

상기 프로세서는,

상기 페이지랭크 알고리즘으로부터 페이지랭크 벡터를 획득하는 벡터 획득 모듈;

상기 페이지랭크 벡터의 유클리드 거리값을 연산하는 거리값 연산 모듈; 및

상기 유클리드 거리값을 기초로 상기 희소행렬 연산의 부동소수점 정밀도를 변환하는 정밀도 변환 모듈;

을 포함하는 연산 장치.

청구항 2

제1항에 있어서,

상기 희소행렬 연산은 상기 페이지랭크 알고리즘의 각 페이지의 인접행렬과 상기 페이지랭크 벡터의 희소행렬 곱셈 연산을 포함하는, 연산 장치.

청구항 3

제1항에 있어서,

상기 거리값 연산 모듈은, 시계열적으로 연속하는 제1 페이지랭크 벡터 및 제2 페이지랭크 벡터 사이의 유클리드 거리값을 연산하는, 연산 장치.

청구항 4

제1항에 있어서,

상기 정밀도 변환 모듈은, 상기 유클리드 거리값이 미리 설정된 전환값을 초과하는 경우 상기 희소행렬 연산의 부동소수점 정밀도를 단정밀도에서 반정밀도로 변환하는, 연산 장치.

청구항 5

제4항에 있어서,

상기 정밀도 변환 모듈은, 상기 유클리드 거리값이 미리 설정된 전환값 이하인 경우 상기 희소행렬 연산의 부동소수점 정밀도를 반정밀도에서 단정밀도로 변환하고,

상기 미리 설정된 전환값은 상기 페이지랭크 알고리즘의 각 페이지의 인접행렬과 상기 페이지랭크 벡터를 기초로 설정되는, 연산 장치.

청구항 6

제1항에 있어서,

상기 프로세서는, 상기 부동소수점 정밀도를 기초로 상기 희소행렬 연산을 수행하는 곱셈연산 모듈을 더 포함하고,

상기 곱셈연산 모듈은, 상기 희소행렬 연산의 부동소수점 정밀도가 단정밀도인 경우 하나의 단정밀도 곱셈 연산을 수행하고, 상기 희소행렬 연산의 부동소수점 정밀도가 반정밀도인 경우 두개의 반정밀도 곱셈 연산을 병렬적으로 수행하는, 연산 장치.

청구항 7

제6항에 있어서,

상기 곱셈연산 모듈은, 상기 하나의 단정밀도 곱셈 연산과 상기 두개의 반정밀도 곱셈 연산에서 가수 부분의 곱셈기를 공유하여 동일한 곱셈기를 이용하여 연산을 수행하는, 연산 장치.

청구항 8

페이지랭크 알고리즘의 희소행렬 연산을 수행하는 연산 방법에 있어서,

상기 페이지랭크 알고리즘으로부터 페이지랭크 벡터를 획득하는 단계;

상기 페이지랭크 벡터의 유클리드 거리값을 연산하는 단계; 및

상기 유클리드 거리값을 기초로 상기 희소행렬 연산의 부동소수점 정밀도를 변환하는 단계;

를 포함하는, 연산 방법.

청구항 9

제8항에 있어서,

상기 희소행렬 연산은 상기 페이지랭크 알고리즘의 각 페이지의 인접행렬과 상기 페이지랭크 벡터의 희소행렬 곱셈 연산을 포함하는, 연산 방법.

청구항 10

제8항에 있어서,

상기 거리값을 연산하는 단계는, 시계열적으로 연속하는 제1 페이지랭크 벡터 및 제2 페이지랭크 벡터 사이의 유클리드 거리값을 연산하는 단계를 포함하는, 연산 방법.

청구항 11

제8항에 있어서,

상기 부동소수점 정밀도를 변환하는 단계는, 상기 유클리드 거리값이 미리 설정된 전환값을 초과하는 경우 상기 희소행렬 연산의 부동소수점 정밀도를 단정밀도에서 반정밀도로 변환하는 단계를 포함하는, 연산 방법.

청구항 12

제11항에 있어서,

상기 부동소수점 정밀도를 변환하는 단계는, 상기 유클리드 거리값이 미리 설정된 전환값 이하인 경우 상기 희소행렬 연산의 부동소수점 정밀도를 반정밀도에서 단정밀도로 변환하는 단계를 포함하고,

상기 미리 설정된 전환값은 상기 페이지랭크 알고리즘의 각 페이지의 인접행렬과 상기 페이지랭크 벡터를 기초로 설정되는, 연산 방법.

청구항 13

제8항에 있어서,

상기 부동소수점 정밀도를 기초로 상기 희소행렬 연산을 수행하는 단계를 더 포함하고,

상기 희소행렬 연산을 수행하는 단계는,

상기 희소행렬 연산의 부동소수점 정밀도가 단정밀도인 경우 하나의 단정밀도 곱셈 연산을 수행하는 단계; 및

상기 희소행렬 연산의 부동소수점 정밀도가 반정밀도인 경우 두개의 반정밀도 곱셈 연산을 병렬적으로 수행하는 단계를 포함하는, 연산 방법.

청구항 14

제13항에 있어서,

상기 희소행렬 연산을 수행하는 단계는, 상기 하나의 단정밀도 곱셈 연산과 상기 두개의 반정밀도 곱셈 연산에서 가수 부분의 곱셈기를 공유하여 동일한 곱셈기를 이용하여 연산을 수행하는 단계를 포함하는, 연산 방법.

청구항 15

컴퓨터를 이용하여 제8항 내지 제14항 중 어느 한 항의 방법을 실행시키기 위하여 기록매체에 저장된 컴퓨터 프로그램.

발명의 설명

기 술 분 야

[0001] 본 발명의 실시예들은, 페이지랭크 희소행렬 곱셈 연산을 위한 연산 장치 및 방법에 관한 것이다.

배 경 기 술

[0002] 페이지랭크(PageRank, PR) 알고리즘은 구글의 래리 페이지(Larry Page)가 1998년 창안하여 현재까지도 웹 검색 분야에서 널리 사용되고 있다. 페이지랭크 알고리즘에서는 각 페이지들의 중요도를 상징하는 페이지랭크 값을 연산하고, 이후 이 값을 크기순으로 정렬해서 중요도 랭킹을 내보낸다. 이 페이지랭크 값은 그 노드가 다른 노드들과 연결된 링크가 많은 경우, 그리고 이미 페이지랭크 값이 높은, 즉 중요도가 높은 노드와 연결이 되어있을 때 값이 커지는 특징이 있다. 페이지랭크 알고리즘은 오래 전에 개발되었지만, 이를 응용한 여러 알고리즘들이 데이터 사이언스나 그래프 프로세싱에서 널리 사용되고 있다.

[0003] 페이지랭크 알고리즘에서 가장 핵심이 되는 연산은 부동소수점으로 표현된 소수 값의 페이지랭크 벡터를 A(Adjacency Matrix, 인접행렬)과 반복적으로 곱하는 부분이다. 여기서 인접행렬 A는 각 페이지들이 다른 페이지들과 연결성이 있을 때 1의 원소값을 가지는 행렬이다. 일반적으로 페이지랭크는 굉장히 많은 페이지들을 대상으로 연산을 진행하고, 각각의 페이지는 소수의 페이지들과 연결되어 있기 때문에 불가피하게 높은 희소도(sparsity)를 가지게 된다. 따라서 인접행렬 A는 희소행렬이 된다. 그렇기 때문에, 페이지랭크 알고리즘에서는 희소행렬과 페이지랭크 벡터를 곱하는 희소행렬 곱셈(SpMV, Sparse Matirx Vector Multiplication)이 반복적으로 수행되고, 이 부분이 페이지랭크 알고리즘 연산에서 병목현상이 나타나는 부분이 된다.

발명의 내용

해결하려는 과제

[0004] 본 발명이 해결하고자 하는 과제는 전술한 문제점을 극복하기 위한 것으로서, 페이지랭크 희소행렬 곱셈 연산을 위한 연산 장치 및 방법을 제공하는 것이다. 그러나 이러한 과제는 예시적인 것으로, 이에 의해 본 발명의 범위가 한정되는 것은 아니다.

과제의 해결 수단

[0005] 상술한 기술적 과제들을 달성하기 위한 기술적 수단으로서, 본 발명의 일 측면에 따른 연산 장치는, 페이지랭크 알고리즘의 희소행렬 연산을 수행하는 적어도 하나의 프로세서를 포함하는 연산 장치에 있어서, 상기 프로세서는, 상기 페이지랭크 알고리즘으로부터 페이지랭크 벡터를 획득하는 벡터 획득 모듈, 상기 페이지랭크 벡터의 유클리드 거리값을 연산하는 거리값 연산 모듈, 및 상기 유클리드 거리값을 기초로 상기 희소행렬 연산의 부동소수점 정밀도를 변환하는 정밀도 변환 모듈을 포함하는 연산 장치가 제공된다.

[0006] 일 예에 따르면, 상기 희소행렬 연산은 상기 페이지랭크 알고리즘의 각 페이지의 인접행렬과 상기 페이지랭크 벡터의 희소행렬 곱셈 연산을 포함할 수 있다.

[0007] 다른 예에 따르면, 상기 거리값 연산 모듈은, 시계열적으로 연속하는 제1 페이지랭크 벡터 및 제2 페이지랭크 벡터 사이의 유클리드 거리값을 연산할 수 있다.

[0008] 또 다른 예에 따르면, 상기 정밀도 변환 모듈은, 상기 유클리드 거리값이 미리 설정된 전환값을 초과하는 경우 상기 희소행렬 연산의 부동소수점 정밀도를 단정밀도에서 반정밀도로 변환할 수 있다.

[0009] 또 다른 예에 따르면, 상기 정밀도 변환 모듈은, 상기 유클리드 거리값이 미리 설정된 전환값 이하인 경우 상기 희소행렬 연산의 부동소수점 정밀도를 반정밀도에서 단정밀도로 변환하고, 상기 미리 설정된 전환값은 상기 페

이지랭크 알고리즘의 각 페이지의 인접행렬과 상기 페이지랭크 벡터를 기초로 설정될 수 있다.

[0010] 또 다른 예에 따르면, 상기 프로세서는, 상기 부동소수점 정밀도를 기초로 상기 희소행렬 연산을 수행하는 곱셈 연산 모듈을 더 포함하고, 상기 곱셈연산 모듈은, 상기 희소행렬 연산의 부동소수점 정밀도가 단정밀도인 경우 하나의 단정밀도 곱셈 연산을 수행하고, 상기 희소행렬 연산의 부동소수점 정밀도가 반정밀도인 경우 두개의 반정밀도 곱셈 연산을 병렬적으로 수행할 수 있다.

[0011] 또 다른 예에 따르면, 상기 곱셈연산 모듈은, 상기 하나의 단정밀도 곱셈 연산과 상기 두개의 반정밀도 곱셈 연산에서 가수 부분의 곱셈기를 공유하여 동일한 곱셈기를 이용하여 연산을 수행할 수 있다.

[0012] 상술한 기술적 과제들을 달성하기 위한 기술적 수단으로서, 본 발명의 일 측면에 따른 연산 방법은 페이지랭크 알고리즘의 희소행렬 연산을 수행하는 연산 방법에 있어서, 상기 페이지랭크 알고리즘으로부터 페이지랭크 벡터를 획득하는 단계, 상기 페이지랭크 벡터의 유클리드 거리값을 연산하는 단계, 및 상기 유클리드 거리값을 기초로 상기 희소행렬 연산의 부동소수점 정밀도를 변환하는 단계를 포함하는, 연산 방법이 제공된다.

[0013] 일 예에 따르면, 상기 희소행렬 연산은 상기 페이지랭크 알고리즘의 각 페이지의 인접행렬과 상기 페이지랭크 벡터의 희소행렬 곱셈 연산을 포함할 수 있다.

[0014] 다른 예에 따르면, 상기 거리값을 연산하는 단계는, 시계열적으로 연속하는 제1 페이지랭크 벡터 및 제2 페이지랭크 벡터 사이의 유클리드 거리값을 연산하는 단계를 포함할 수 있다.

[0015] 또 다른 예에 따르면, 상기 부동소수점 정밀도를 변환하는 단계는, 상기 유클리드 거리값이 미리 설정된 전환값을 초과하는 경우 상기 희소행렬 연산의 부동소수점 정밀도를 단정밀도에서 반정밀도로 변환하는 단계를 포함할 수 있다.

[0016] 또 다른 예에 따르면, 상기 부동소수점 정밀도를 변환하는 단계는, 상기 유클리드 거리값이 미리 설정된 전환값 이하인 경우 상기 희소행렬 연산의 부동소수점 정밀도를 반정밀도에서 단정밀도로 변환하는 단계를 포함하고, 상기 미리 설정된 전환값은 상기 페이지랭크 알고리즘의 각 페이지의 인접행렬과 상기 페이지랭크 벡터를 기초로 설정될 수 있다.

[0017] 또 다른 예에 따르면, 상기 부동소수점 정밀도를 기초로 상기 희소행렬 연산을 수행하는 단계를 더 포함하고, 상기 희소행렬 연산을 수행하는 단계는, 상기 희소행렬 연산의 부동소수점 정밀도가 단정밀도인 경우 하나의 단정밀도 곱셈 연산을 수행하는 단계, 및 상기 희소행렬 연산의 부동소수점 정밀도가 반정밀도인 경우 두개의 반정밀도 곱셈 연산을 병렬적으로 수행하는 단계를 포함할 수 있다.

[0018] 또 다른 예에 따르면, 상기 희소행렬 연산을 수행하는 단계는, 상기 하나의 단정밀도 곱셈 연산과 상기 두개의 반정밀도 곱셈 연산에서 가수 부분의 곱셈기를 공유하여 동일한 곱셈기를 이용하여 연산을 수행하는 단계를 포함할 수 있다.

[0019] 상술한 기술적 과제들을 달성하기 위한 기술적 수단으로서, 컴퓨터를 이용하여 상술한 방법을 실행시키기 위하여 기록매체에 저장된 컴퓨터 프로그램이 제공된다.

[0020] 전술한 것 외의 다른 측면, 특징, 이점은 이하의 발명을 실시하기 위한 구체적인 내용, 청구범위 및 도면으로부터 명확해질 것이다.

발명의 효과

[0021] 상기한 바와 같이 이루어진 본 발명의 일 실시예에 따르면, 희소행렬 곱셈을 연산하기 위한 부동소수점 정밀도 변환 알고리즘을 통해 효과적으로 페이지랭크 연산을 수행할 수 있는 페이지랭크 희소행렬 곱셈 연산을 위한 연산 장치 및 방법을 구현할 수 있다. 물론 이러한 효과에 의해 본 발명의 범위가 한정되는 것은 아니다.

도면의 간단한 설명

[0022] 도 1은 본 발명의 일 실시예에 따른 연산 방법을 개략적으로 설명하는 순서도이다.

도 2는 본 발명의 일 실시예에 따른 연산 장치가 포함하는 프로세서의 구성을 개략적으로 도시하는 도면이다.

도 3은 본 발명의 일 실시예에 따른 연산 방법을 설명하기 위한 순서도이다.

도 4 및 도 5는 본 발명의 일 실시예에 따른 연산 장치가 수행하는 페이지랭크 희소행렬 곱셈 연산 방법을 설명

하기 위한 도면이다.

도 6 및 도 7은 본 발명의 일 실시예에 따른 연산 장치의 성능을 실험한 실험 결과의 예시들이다.

발명을 실시하기 위한 구체적인 내용

[0023] 아래에서는 첨부한 도면을 참조하여 본 개시가 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 다양한 실시예들을 상세히 설명한다. 그러나 본 개시의 기술적 사상은 다양한 형태로 변형되어 구현될 수 있으므로 본 명세서에서 설명하는 실시예들로 제한되지 않는다. 본 명세서에 개시된 실시예들을 설명함에 있어서 관련된 공지 기술을 구체적으로 설명하는 것이 본 개시의 기술적 사상의 요지를 흐릴 수 있다고 판단되는 경우 그 공지 기술에 대한 구체적인 설명을 생략한다. 동일하거나 유사한 구성요소는 동일한 참조 번호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다.

[0024] 본 명세서에서 어떤 요소가 다른 요소와 "연결"되어 있다고 기술될 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라 그 중간에 다른 요소를 사이에 두고 "간접적으로 연결"되어 있는 경우도 포함한다. 어떤 요소가 다른 요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 요소 외에 또 다른 요소를 배제하는 것이 아니라 또 다른 요소를 더 포함할 수 있는 것을 의미한다.

[0025] 일부 실시예들은 기능적인 블록 구성들 및 다양한 처리 단계들로 설명될 수 있다. 이러한 기능 블록들의 일부 또는 전부는 특정 기능을 실행하는 다양한 개수의 하드웨어 및/또는 소프트웨어 구성들로 구현될 수 있다. 예를 들어, 본 개시의 기능 블록들은 하나 이상의 마이크로프로세서들에 의해 구현되거나, 소정의 기능을 위한 회로 구성들에 의해 구현될 수 있다. 본 개시의 기능 블록들은 다양한 프로그래밍 또는 스크립팅 언어로 구현될 수 있다. 본 개시의 기능 블록들은 하나 이상의 프로세서들에서 실행되는 알고리즘으로 구현될 수 있다. 본 개시의 기능 블록이 수행하는 기능은 복수의 기능 블록에 의해 수행되거나, 본 개시에서 복수의 기능 블록이 수행하는 기능들은 하나의 기능 블록에 의해 수행될 수도 있다. 또한, 본 개시는 전자적인 환경 설정, 신호 처리, 및/또는 데이터 처리 등을 위하여 종래 기술을 채용할 수 있다.

[0026] 도 1은 본 발명의 일 실시예에 따른 연산 방법을 개략적으로 설명하는 순서도이고, 도 2는 본 발명의 일 실시예에 따른 연산 장치가 포함하는 프로세서의 구성을 개략적으로 도시하는 도면이다.

[0027] 먼저, 도 2를 참조하면, 본 발명의 일 실시예에 따른 연산 장치는 적어도 하나의 프로세서(110)를 포함할 수 있다. 하지만 본 발명은 이에 한정되지 않으며, 연산 장치는 다른 구성요소를 더 포함할 수도 있고 일부 구성요소가 생략될 수도 있다. 연산 장치의 일부 구성요소는 복수의 장치로 분리될 수도 있고, 복수개의 구성요소들이 하나의 장치로 병합될 수도 있다.

[0028] 본 발명의 일 실시예에 따른 프로세서(110)는 메모리를 포함할 수도 있다. 하지만 본 발명은 이에 한정되지 않으며, 프로세서(110) 및 메모리는 연산 장치에 포함될 수 있다.

[0029] 프로세서(110)는 메모리에 저장된 데이터의 검색, 추출, 연산, 읽기, 쓰기 등의 동작을 수행할 수 있다. 프로세서(110)는 마이크로프로세서(Microprocessor), 중앙처리장치(Central Processing Unit: CPU), 프로세서 코어(Processor Core), 멀티프로세서(Multiprocessor), ASIC(Application-Specific Integrated Circuit), FPGA(Field Programmable Gate Array) 등의 처리 장치를 포함할 수 있으나, 본 발명은 이에 한정되지 않는다.

[0030] 메모리는 컴퓨터에서 판독 가능한 기록 매체로서, RAM(random access memory), ROM(read only memory) 및 디스크 드라이브와 같은 비소멸성 대용량 기록장치(permanent mass storage device)를 포함할 수 있다. 또한, 메모리에는 연산 장치를 제어하기 위한 프로그램 코드 및 데이터가 일시적 또는 영구적으로 저장될 수 있다.

[0031] 도 1 및 도 2에 도시된 바와 같이, 본 발명의 일 실시예에 따른 연산 장치는 적어도 하나의 프로세서(110)를 이용하여 도 1의 연산 방법을 수행할 수 있다. 예를 들어, 도 1에 도시된 바와 같이, 프로세서(110)는 페이지랭크 알고리즘의 희소행렬 연산을 수행할 수 있다.

[0032] 도 1을 참조하면, 단계 S100에서, 본 발명의 일 실시예에 따른 프로세서(110)는 페이지랭크 벡터의 유클리드 거리값을 연산하여 차이값(DIFF)으로 정의하고, 차이값(DIFF)을 기초로 희소행렬 연산의 부동소수점 정밀도를 변환(Transprecision)할 수 있다. 구체적으로, 차이값(DIFF)은 기존의 페이지랭크 벡터값과 최신의 페이지랭크 벡터값 사이의 유클리드 거리값(Euclidean distance)일 수 있다.

[0033] 단계 S150에서, 프로세서(110)는 차이값(DIFF)이 미리 설정된 전환값(TP)을 초과하는 경우 반정밀도(Half precision)로 부동 소수점 정밀도를 전환하여 희소 행렬 곱셈 연산을 수행한다.

[0034] 단계 S200 및 S250에서, 프로세서(110)는 차이값(DIFF)이 미리 설정된 전환값(TP) 이하인 경우 단정밀도(Single precision)로 부동 소수점 정밀도를 전환하여 희소 행렬 곱셈 연산을 수행한다. 또한, 프로세서(110)는 차이값(DIFF)이 미리 설정된 임계값(ε)에 해당하는 경우 희소 행렬 곱셈 연산을 종료한다. 예컨대, 미리 설정된 임계값(ε)은 페이지랭크 알고리즘의 연산을 종료하는 임계값을 나타낼 수 있다.

[0035] 본 발명의 일 실시예에 따른 연산 장치는 페이지랭크 알고리즘의 희소 행렬 곱셈 연산에서 부동 소수점 계산 정밀도를 조건부로 변경하여 연산의 성능을 향상 시킬 수 있다.

[0036] 도 2를 참조하면, 프로세서(110)는 벡터 획득 모듈(111), 거리값 연산 모듈(113), 정밀도 변환 모듈(115), 및 곱셈 연산 모듈(117)을 포함할 수 있다. 프로세서(110)와 이에 포함된 모듈들은, 전적으로 하드웨어일 수 있고, 또는 부분적으로 하드웨어이고 부분적으로 소프트웨어인 측면을 가질 수 있다. 예컨대, 프로세서(110)와 이에 포함된 각 모듈은 특정 형식 및 내용의 데이터를 처리하거나 전자통신 방식으로 주고받기 위한 하드웨어 및 이에 관련된 소프트웨어를 통칭할 수 있다. 본 명세서에서 "모듈"의 용어는 하드웨어 및 해당 하드웨어에 의해 구동되는 소프트웨어의 조합을 지칭하는 것으로 해석될 수 있다. 예컨대, 하드웨어는 CPU 또는 다른 프로세서(processor)를 포함하는 데이터 처리 기기일 수 있다. 또한, 하드웨어에 의해 구동되는 소프트웨어는 실행중인 프로세스, 객체(object), 실행파일(executable file), 실행 스레드(thread of execution), 프로그램(program) 등을 지칭할 수 있다.

[0037] 또한, 프로세서(110)를 구성하는 각각의 요소는 반드시 서로 물리적으로 구분되는 별개의 장치를 지칭하는 것으로 한정되는 것은 아니다. 즉, 도 2의 벡터 획득 모듈(111), 거리값 연산 모듈(113), 정밀도 변환 모듈(115), 및 곱셈 연산 모듈(117) 등은 프로세서(110)에 의해 수행되는 동작에 따라 기능적으로 구분한 것일 뿐, 반드시 서로 독립적으로 구분되어야 하는 것이 아니다. 물론, 실시예에 따라서는 이들은 서로 물리적으로 구분되는 별개의 장치로 구현되는 것도 가능하다.

[0038] 이하에서는 상술한 것과 같은 구성을 취할 수 있는 프로세서(110)의 각 모듈들의 동작 등에 대해, 연산 방법과 함께 설명한다.

[0039] 도 3은 본 발명의 일 실시예에 따른 연산 방법을 설명하기 위한 순서도이다.

[0040] 도 4 및 도 5는 본 발명의 일 실시예에 따른 연산 장치가 수행하는 페이지랭크 희소행렬 곱셈 연산 방법을 설명하기 위한 도면이다.

[0041] 먼저, 도 3을 참조하면, 단계 S310에서, 벡터 획득 모듈(111)은 페이지랭크 알고리즘으로부터 페이지랭크 벡터를 획득할 수 있다.

[0042] 단계 S320에서, 거리값 연산 모듈(113)은 페이지랭크 벡터의 유클리드 거리값을 연산할 수 있다. 예를 들어, 거리값 연산 모듈(113)은, 시계열적으로 연속하는 제1 페이지랭크 벡터 및 제2 페이지랭크 벡터 사이의 유클리드 거리값을 연산할 수 있다.

[0043] 단계 S330에서, 정밀도 변환 모듈(115)은 유클리드 거리값을 기초로 희소행렬 연산의 부동소수점 정밀도를 변환할 수 있다. 또한, 곱셈 연산 모듈(117)은 희소행렬 연산을 수행할 수 있다. 예를 들어, 곱셈 연산 모듈(117)은 페이지랭크 알고리즘의 각 페이지의 인접행렬과 페이지랭크 벡터의 희소행렬 곱셈 연산을 수행할 수 있다.

[0044] 도 3 내지 도 5를 함께 참조하면, 본 발명의 일 실시예에 따른 연산 장치는 2개의 SIMD(Single Instruction Multiple Data) FP16 장치(S425) 또는 1개의 FP32 장치(S435)를 사용할 수 있는 다중 모드 부동 소수점 SpMV 엔진을 포함할 수 있다. 예컨대, 도 4에 도시된 바와 같이, 반정밀도 작동 모드에서는 2개의 FP16 SpMV 장치(S425)가 병렬로 작동하여 전체 연산 실행 시간의 소비를 절반으로 줄일 수 있다.

[0045] 예를 들어, 도 4에 도시된 바와 같이, 페이지랭크의 데이터가 Coordinate(COO) 형식으로 저장될 수 있다(S410). 이어서, 본 발명의 일 실시예에 따른 연산 장치는 반정밀도(FP16) 모드에서 작동할 수 있다. 이어서, 단계 S420에서, 차이값(DIFF)이 전환값(TP)보다 낮아지면, 본 발명의 일 실시예에 따른 연산 장치는 차이값(DIFF)이 임계값(ε)에 도달할 때까지 계산 정밀도를 단정밀도(FP32) 모드로 변환할 수 있다. 기존의 페이지랭크 알고리즘과 마찬가지로 차이값(DIFF)이 임계값(ε)으로 수렴하면 전체 연산이 종료될 수 있다(S430).

[0046] 정밀도 변환 모듈(115)은, 유클리드 거리값이 미리 설정된 전환값을 초과하는 경우 희소행렬 연산의 부동소수점 정밀도를 단정밀도에서 반정밀도로 변환할 수 있다.

[0047] 정밀도 변환 모듈(115)은, 유클리드 거리값이 미리 설정된 전환값 이하인 경우 희소행렬 연산의 부동소수점 정

밀도를 반정밀도에서 단정밀도로 변환할 수 있다. 예컨대, 미리 설정된 전환값은 페이지랭크 알고리즘의 각 페이지의 인접행렬과 페이지랭크 벡터를 기초로 설정될 수 있다.

[0048] 본 발명에 따른 연산 장치는, 이와 같은 부동소수점 정밀도 변환 알고리즘을 효율적으로 수행하기 위한 희소행렬 곱셈 연산기 하드웨어 구조도 함께 제시하였다. 본 발명의 일 실시예에 따른 연산 장치에서는 하나의 단정밀도 곱셈을 수행하는 모드와 두 개의 반정밀도 곱셈을 동시에 수행하는 두 가지의 모드가 존재한다. 페이지랭크 연산이 시작되면 차이값이 특정 전환값에 도달할 때까지 두 개의 반정밀도 곱셈이 Single Instruction Multiple Data, SIMD로 동시에 수행된다. 병렬적으로 두 개의 곱셈을 동시에 수행할 수 있기 때문에, 하나씩 연산할 때보다 수행시간이 절반으로 줄어든다는 장점이 있다. 차이값이 전환값보다 작아지게 되면, 하나의 단정밀도 곱셈이 수행되도록 연산기가 동작한다. 단정밀도는 반정밀도보다 bit의 너비가 크지만 정밀하다는 장점이 있기 때문에 연산 속도는 느려지지만 좀더 정밀하게 연산이 수행된다.

[0049] 본 발명에 따른 연산 장치에서 이렇게 두 가지의 모드를 효율적으로 지원하기 위해서 단정밀도와 반정밀도 연산에서 부동소수점 숫자의 가수(mantissa) 부분의 곱셈기를 공유할 수 있다. 부동소수점 숫자의 곱셈에서 가장 큰 자원을 소모하는 부분이 가수의 곱셈 부분이기 때문에, 이 부분을 공유함으로써 자원을 절약할 수 있다. 이를 위하여 나누어 정복하는(divide and conquer) 알고리즘을 사용하여 두 모드 사이에서 같은 곱셈기를 사용할 수 있다.

[0050] 일반적으로 부동 소수점 곱셈 연산에서 두 피연산자의 지수 부분은 덧셈기를 사용하여 간단하게 계산할 수 있다. 그러나 가수 비트는 곱해야 하며 이는 부동 소수점 곱셈의 막대한 오버헤드이다. 이 문제를 해결하기 위해 본 발명에서는 가수 비트의 승수 단위를 FP16과 FP32가 공유하는 하드웨어 아키텍처를 제안했다.

[0051] 예를 들어, 도 5를 참조하면, 본 발명의 일 실시예에 따른 32비트 단정밀도에서 작동할 때 연산 장치의 아키텍처가 도시되어 있다. 여기서, X 및 Y는 두 단정밀도 숫자의 유효숫자(significand)를 나타내므로 각각은 24비트(n=24)를 가질 수 있다. 그리고, X 및 Y는 12비트로 분리될 수 있다. 여기서, XL 및 YL은 낮은 부분이고, XH 및 YH는 높은 부분이다.

[0052] 12비트 곱셈기(510, 520)는 병렬로 작동하므로 SIMD FP16 반정밀 작동 모드에서도 사용할 수 있다. 반정밀도 유형의 가수는 10비트이므로 이러한 승수는 두 개의 반정밀도 SpMV 계산을 위한 승수로 작동할 수 있다.

[0053] <수학식1>

$$X = X_H \cdot 2^{n/2} + X_L$$

$$Y = Y_H \cdot 2^{n/2} + Y_L$$

$$X \cdot Y = (X_H \cdot 2^{n/2} + X_L) \cdot (Y_H \cdot 2^{n/2} + Y_L)$$

$$= X_H \cdot Y_H \cdot 2^n + X_H \cdot Y_L \cdot 2^{n/2} + X_L \cdot Y_H \cdot 2^{n/2} + X_L \cdot Y_L$$

[0054]

[0055] 단정밀도 FP32 모드에서 <수학식1>과 같이 분할된 유효 피연산자는 FP16 12비트 승수에 들어간 다음 시프터로 전달될 수 있다. $X_H \cdot Y_L$ 및 $X_L \cdot Y_H$ 은 12비트 시프터를 한 번 통과해야 하고, $X_H \cdot Y_H$ 는 두 번 통과해야 하며, $X_L \cdot Y_L$ 은 시프터를 통과할 필요가 없다. 반정밀도 FP16 작동 모드에서는 시프터 없이 병렬로 2개의 12비트 승수를 사용하여 SIMD에서 10비트 가수를 계산할 수 있다. 승수 및 시프터로 전달된 데이터는 부동 소수점 정밀도의 정규화 및 반올림 논리 회로로 전달될 수 있다.

[0056] 본 발명의 일 실시예에 따른 연산 장치에 따르면, 2개의 12비트 곱셈기를 이용하여, FP16 및 FP32 작동 모드 간에 승수를 공유하여 리소스를 절약할 수 있다.

[0057] 본 발명의 일 실시예에 따른 프로세서는, 부동소수점 정밀도를 기초로 희소행렬 연산을 수행하는 곱셈연산 모듈(117)을 더 포함할 수 있다.

[0058] 곱셈연산 모듈(117)은, 희소행렬 연산의 부동소수점 정밀도가 단정밀도인 경우 하나의 단정밀도 곱셈 연산을 수행하고, 희소행렬 연산의 부동소수점 정밀도가 반정밀도인 경우 두개의 반정밀도 곱셈 연산을 병렬적으로 수행할 수 있다.

[0059] 곱셈연산 모듈(117)은, 하나의 단정밀도 곱셈 연산과 두개의 반정밀도 곱셈 연산에서 가수 부분의 곱셈기를 공유하여 동일한 곱셈기를 이용하여 연산을 수행할 수 있다.

[0060] 도 6 및 도 7은 본 발명의 일 실시예에 따른 연산 장치의 성능을 실험한 실험 결과의 예시들이다.

[0061] 도 6 및 도 7을 참조하면, 실험에 사용된 데이터 세트와 실험 결과가 도시되어 있다. 예를 들어, 도 6에 도시된 바와 같이, Twitter, Stanford, Amazon 데이터셋은 소셜네트워크, 웹그래프, 전자상거래 네트워크의 각 분야에서 선정되었으며, 도 6은 각 데이터 세트의 전체 정보를 나타낸다.

[0062] 도 7을 참조하면, 각 데이터 세트의 오류 수가 도시되어 있다. 여기서, 'half'는 반정밀도(Half precision)를 나타내고, 'single'은 단정밀도(Single precision)를 나타내고, 'trans'는 본 발명에 따른 정밀도 전환(Half precision에서 Single precision 으로 전환)의 실시예를 나타낸다.

[0063] 예를 들어, 세 가지 유형의 정밀도는 모두 10^{-5}로 설정된 임계값으로 차이값(DIFF)이 수렴할 때까지 페이지랭크 연산을 수행할 수 있다. 정밀도 유형 'trans'에서는 10^{-3} 전환값(TP)을 사용하여 처음에 반정밀도로 연산하다가 단정밀도로 변환하여 페이지랭크 연산을 수행할 수 있다. 오류의 수는 가장 정확한 계산 결과인 배정밀도(Double precision)로 연산된 상위 50/100 순위 결과와 다른 순위의 수를 세어 결정할 수 있다.

[0064] 반정밀도는 단정밀도보다 적은 리소스로 더 빠른 속도로 페이지랭크 연산을 수행할 수 있지만 단정밀도에 비해 오류가 많다. 그러나, 도 7에 도시된 바와 같이, 본 발명에 따른 정밀도 전환에서는 top50 및 top100 순위 모두에서 단정밀도의 결과와 동일한 오류 수가 발생하는 것을 알 수 있다. 즉, 페이지랭크 연산을 수행하는 동안 부동 소수점의 정밀도 유형을 변경하면 반정밀도 계산에서 발생할 수 있는 많은 수의 오류를 극복하면서 속도의 이점을 얻을 수 있다.

[0065] 결국, 본 발명에 따른 연산 방법을 사용하면 페이지랭크 연산이 반정밀도보다 정확하고 단정밀도보다 빠른 효과가 있다. 또한, 본 발명에 따른 연산 장치는 하드웨어 리소스를 공유함으로써 최소한의 하드웨어 오버헤드로 기존 단정밀도 연산기보다 빠르게 연산을 수행할 수 있다.

[0066] 본 발명에 따르면, 부동소수점 정밀도를 변환하는 과정을 통해 기존에 단정밀도만으로 연산할 때 보다는 더욱 빠르게 연산할 수 있고, 반정밀도만으로 연산할 때 보다는 더욱 정밀하게 연산할 수 있다는 특징이 있다.

[0067] 희소행렬 곱셈기는 딥러닝, 그래프 처리 등의 다양한 분야에서 활용되고 있는 연산기이다. 특히 현재까지도 널리 사용되고 있는 웹 검색 알고리즘들의 기본이 되는 페이지랭크 연산에서 가장 큰 부분을 차지하고 있다.

[0068] 본 발명에서는 페이지랭크에서 희소행렬 곱셈을 효율적으로 연산할 수 있는 부동소수점 정밀도 변환 알고리즘을 창안하고, 이를 하드웨어 자원을 아끼면서 처리할 수 있는 하드웨어 구조를 함께 제시하였다. 시뮬레이터를 사용하여 검증한 결과, 연산 중간에 부동소수점 정밀도를 단정밀도에서 반정밀도로 바꾸어 기존의 단정밀도 연산보다 빠르게 처리할 수 있었음에도 정확도 면에서는 단정밀도와 큰 차이를 보이지 않았다. 또한 반정밀도로만 연산했을 때와 비교해서는 훨씬 더 높은 정확도를 보였다.

[0069] 따라서 본 발명의 희소행렬 곱셈기는 연산 중간에 부동소수점의 정밀도를 변환하는 과정을 통해 기존의 단정밀도보다는 빠르게, 반정밀도보다는 정확하게 연산할 수 있는 장점을 가지고 있다. 그리고 본 발명의 하드웨어 연산기 구조는 가수 곱셈기를 단정밀도 연산 모드와 반정밀도 연산 모드에서 공유할 수 있도록 하여 이 알고리즘을 수행할 때 소요되는 하드웨어 자원도 절약하도록 하였다.

[0070] 이상 설명된 다양한 실시예들은 예시적이며, 서로 구별되어 독립적으로 실시되어야 하는 것은 아니다. 본 명세서에서 설명된 실시예들은 서로 조합된 형태로 실시될 수 있다.

[0071] 이상 설명된 다양한 실시예들은 컴퓨터 상에서 다양한 구성요소를 통하여 실행될 수 있는 컴퓨터 프로그램의 형태로 구현될 수 있으며, 이와 같은 컴퓨터 프로그램은 컴퓨터로 판독 가능한 매체에 기록될 수 있다. 이때, 매체는 컴퓨터로 실행 가능한 프로그램을 계속 저장하거나, 실행 또는 다운로드를 위해 임시 저장하는 것일 수도 있다. 또한, 매체는 단일 또는 수개 하드웨어가 결합된 형태의 다양한 기록수단 또는 저장수단일 수 있는데, 어떤 컴퓨터 시스템에 직접 접속되는 매체에 한정되지 않고, 네트워크 상에 분산 존재하는 것일 수도 있다. 매체의 예시로는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium), 및 ROM, RAM, 플래시 메모리 등을 포함하여 프로그램 명령어가 저장되도록 구성된 것이 있을 수 있다. 또한, 다른 매체의 예시로, 애플리케이션을 유통하는 앱 스토어나 기타 다양한 소프트웨어를 공급 내지 유통하는 사이트, 서버 등에서 관리하는

기록매체 내지 저장매체도 들 수 있다.

[0072] 본 명세서에서, "부", "모듈" 등은 프로세서 또는 회로와 같은 하드웨어 구성(hardware component), 및/또는 프로세서와 같은 하드웨어 구성에 의해 실행되는 소프트웨어 구성(software component)일 수 있다. 예를 들면, "부", "모듈" 등은 소프트웨어 구성 요소들, 객체 지향 소프트웨어 구성 요소들, 클래스 구성 요소들 및 태스크 구성 요소들과 같은 구성 요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로 코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들 및 변수들에 의해 구현될 수 있다.

[0073] 전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.

[0074] 본 발명의 범위는 상기 상세한 설명보다는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

부호의 설명

[0075] 110: 프로세서

111: 벡터 획득 모듈

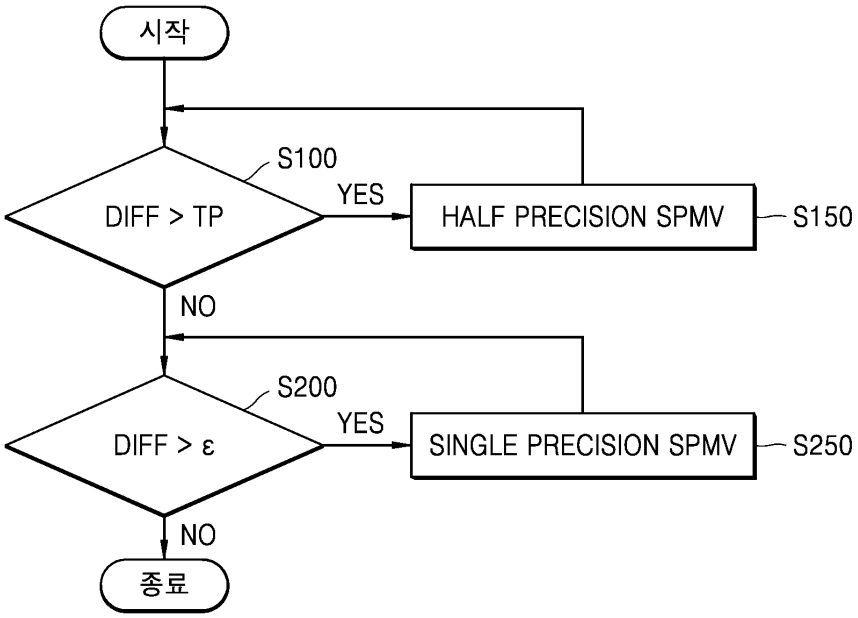
113: 거리값 연산 모듈

115: 정밀도 변환 모듈

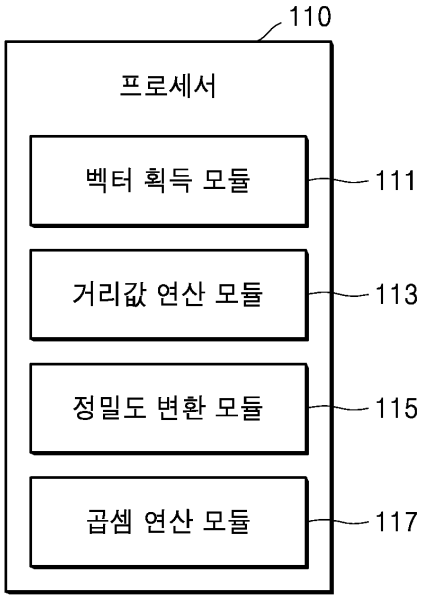
117: 곱셈 연산 모듈

도면

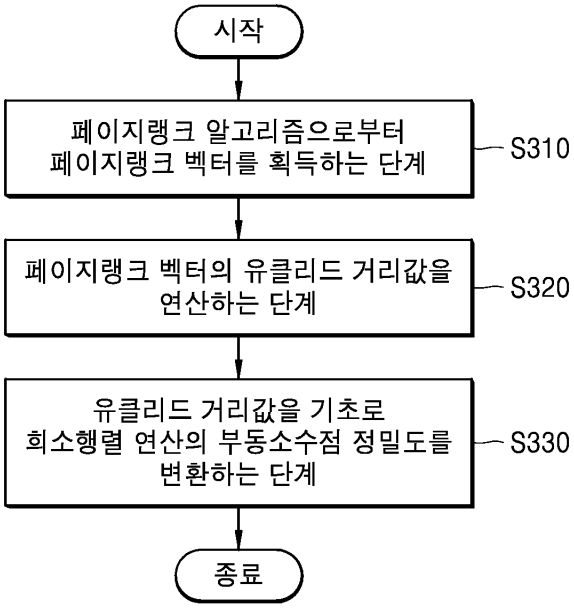
도면1



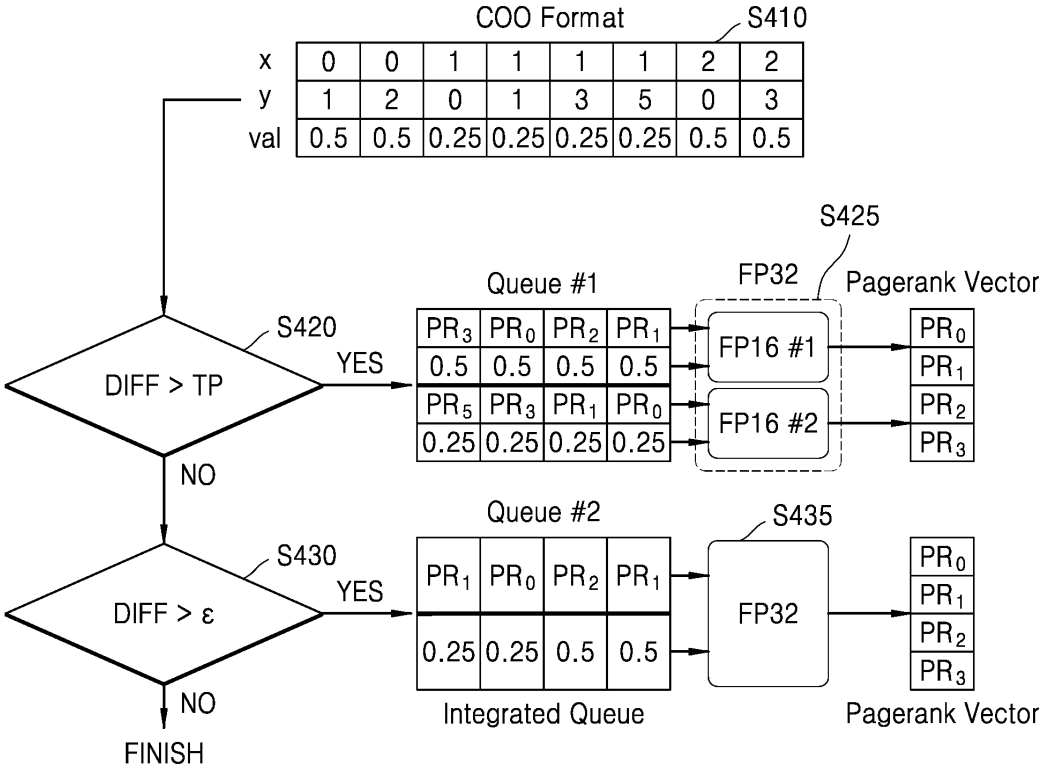
도면2


110
프로세서
벡터 획득 모듈
111
거리값 연산 모듈
113
정밀도 변환 모듈
115
곱셈 연산 모듈
117

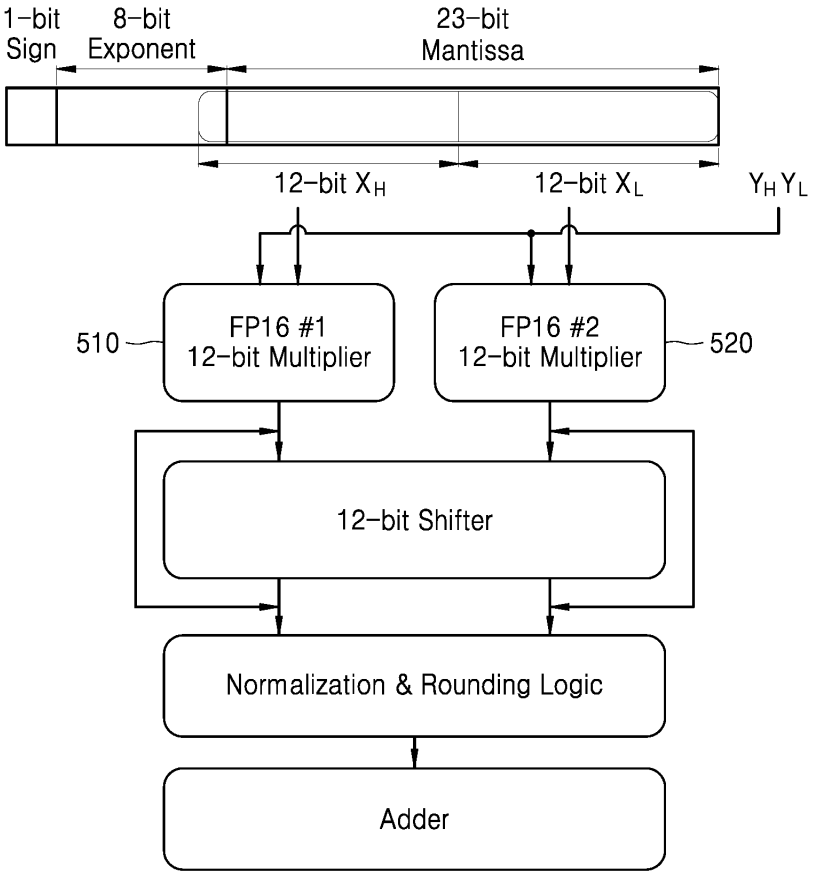
도면3


시작
페이지랭크 알고리즘으로부터
페이지랭크 벡터를 획득하는 단계
S310
페이지랭크 벡터의 유클리드 거리값을
연산하는 단계
S320
유클리드 거리값을 기초로
희소행렬 연산의 부동소수점 정밀도를
변환하는 단계
S330
종료

도면4


COO Format
S410
x 0 0 1 1 1 1 2 2
y 1 2 0 1 3 5 0 3
val 0.5 0.5 0.25 0.25 0.25 0.25 0.5 0.5
S425
FP32
Pagerank Vector
Queue #1
S420
YES
DIFF > TP
PR3 PR0 PR2 PR1
0.5 0.5 0.5 0.5
PR5 PR3 PR1 PR0
0.25 0.25 0.25 0.25
FP16 #1
FP16 #2
PR0 PR1 PR2 PR3
NO
Queue #2
S435
S430
YES
DIFF > ε
PR1 PR0 PR2 PR1
0.25 0.25 0.5 0.5
FP32
PR0 PR1 PR2 PR3
NO
FINISH
Integrated Queue
Pagerank Vector

도면5


1-bit Sign
8-bit Exponent
23-bit Mantissa
12-bit XH
12-bit XL
YH YL
510
FP16 #1 12-bit Multiplier
FP16 #2 12-bit Multiplier
520
12-bit Shifter
Normalization & Rounding Logic
Adder

도면6

Dataset	Type	#of Nodes	#of Edges
Twitter	Social Network	81,306	1,768,149
Stanford	Web Graph	281,903	2,312,497
Amazon	E-Commerce	262,111	1,234,877

도면7

