(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0104279
(43) 공개일자 2025년07월08일

(51) 국제특허분류(Int. Cl.)
G06N 3/063 (2023.01) *G06F 7/483* (2006.01)
G06F 7/544 (2017.01) *G06N 3/0495* (2023.01)
(52) CPC특허분류
G06N 3/063 (2013.01)
G06F 7/483 (2021.08)
(21) 출원번호 10-2023-0196506
(22) 출원일자 2023년12월29일
심사청구일자 없음

(71) 출원인
삼성전자주식회사
경기도 수원시 영통구 삼성로 129 (매탄동)
울산과학기술원
울산광역시 울주군 언양읍 유니스트길 50
(72) 발명자
심현욱
경기도 용인시 기흥구 흥덕1로62번길 7, 201호 (영덕동)
손문규
경기도 용인시 기흥구 서그내로16번길 30, 104동 1104호 (서천동, 영통로 효성해링턴플레이스)
(뒷면에 계속)
(74) 대리인
특허법인 무한

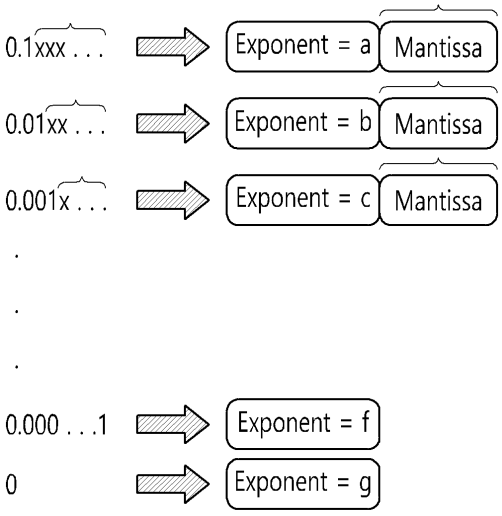
전체 청구항 수 : 총 17 항

(54) 발명의 명칭 **파라미터를 가변적으로 표현하는 방법 및 이를 위한 프로세서 장치**

(57) 요 약

실시예는, 파라미터를 가변적으로 표현하는 방법 및 이를 위한 장치에 관한 것이다. 실시예에 뉴럴 프로세서 장치는, 변환된 파라미터 값의 고정된 길이의 지수부를 읽어오고 상기 지수부의 값에 매핑된 가수부의 길이 정보를 획득하는 비교기; 상기 변환된 파라미터 값의 가수부를 읽어 오고 상기 가수부의 길이 정보를 이용하여 상기 변환된 파라미터 값의 구조를 변환하는 변환기; 및 상기 지수부의 값에 따른 상기 가수부의 길이 정보가 매핑되는 매핑 테이블을 포함할 수 있다.

대 표 도 - 도1



(52) CPC특허분류
G06F 7/5443 (2013.01)
G06N 3/0495 (2023.01)
(72) 발명자
이수길
울산광역시 울주군 언양읍 유니스트길 50
이종은
울산광역시 울주군 언양읍 유니스트길 50
홍민욱
울산광역시 울주군 언양읍 유니스트길 50

명 세 서

청구범위

청구항 1

뉴럴 프로세서 장치에 있어서,

변환된 파라미터 값의 고정된 길이의 지수부를 읽어오고 상기 지수부의 값에 매핑된 가수부의 길이 정보를 획득하는 비교기;

상기 변환된 파라미터 값의 가수부를 읽어 오고 상기 가수부의 길이 정보를 이용하여 상기 변환된 파라미터 값의 구조를 변환하는 변환기; 및

상기 지수부의 값에 따른 상기 가수부의 길이 정보가 매핑되는 매핑 테이블

을 포함하는,

뉴럴 프로세서 장치.

청구항 2

제1항에 있어서,

상기 매핑 테이블은,

파라미터들의 분포에 따라 상기 지수부의 값에 따른 상기 가수부의 길이 정보가 결정되는,

뉴럴 프로세서 장치.

청구항 3

제1항에 있어서,

상기 변환기는,

상기 가수부의 길이 정보에 기초하여 상기 가수부의 값의 자리를 이동시키는,

뉴럴 프로세서 장치.

청구항 4

제1항에 있어서,

상기 변환기는,

상기 변환된 파라미터 값을 표현하기 위해 할당되는 비트 수에 맞추어 0을 추가하는,

뉴럴 프로세서 장치.

청구항 5

제1항에 있어서,

상기 변환된 파라미터 값의 표현 방식에 기초하여,

상기 비교기는, 상기 파라미터 값의 고정된 길이의 실수부의 값에 매핑된 상기 실수부의 지수 이동 정보를 획득

하고,

상기 변환기는, 상기 실수부의 지수 이동 정보에 따라 상기 실수부의 값을 이동하는,

뉴럴 프로세서 장치.

청구항 6

제5항에 있어서,

상기 변환기는,

상기 파라미터 값의 소수부의 값 및 상기 실수부의 지수 이동 정보만큼 이동된 상기 실수부의 값을 합산하는,

뉴럴 프로세서 장치.

청구항 7

제1항에 있어서,

상기 비교기 및 상기 변환기는, SRAM으로부터 상기 변환된 파라미터의 지수부 및 가수부를 각각 불러오는,

뉴럴 프로세서 장치.

청구항 8

제1항에 있어서,

MAC 연산을 위해, 상기 구조가 변환된 상기 파라미터 값 및 상기 지수부의 값이 연산기로 입력되는,

뉴럴 프로세서 장치.

청구항 9

프로세서 장치의 동작 방법에 있어서,

학습된 모델의 파라미터들을 읽어오는 단계;

상기 파라미터들의 분포에 기초하여 고정 길이의 지수부에 대응하는 가수부의 길이 정보를 포함하는 매핑 테이블을 결정하는 단계; 및

상기 파라미터들의 값들을 고정 길이의 지수부 값 및 가변 길이의 가수부를 포함하는 숫자 포맷으로 변환하는 단계

를 포함하는,

프로세서 장치의 동작 방법.

청구항 10

제9항에 있어서,

상기 매핑 테이블은,

상기 파라미터들의 분포에 따라 상기 지수부에 따른 상기 가수부의 길이 정보가 결정되는,

프로세서 장치의 동작 방법.

청구항 11

제9항에 있어서,

상기 숫자 포맷으로 변환된 지수부 및 가수부를 메모리에 기록하는 단계

를 더 포함하는,

프로세서 장치의 동작 방법.

청구항 12

제9항에 있어서,

상기 파라미터들의 값들을 고정 길이의 지수부 값 및 가변 길이의 가수부를 포함하는 숫자 포맷으로 변환하는 단계는,

상기 가변 길이의 가수부에서 첫 번째 나타나는 1을 제외하여 상기 숫자 포맷으로 변환하는 단계

를 포함하는,

프로세서 장치의 동작 방법.

청구항 13

하드웨어와 결합되어 제9항 내지 제12항 중 어느 하나의 항의 방법을 실행시키기 위하여 컴퓨터 판독 가능 기록 매체에 저장된 컴퓨터 프로그램.

청구항 14

프로세서 장치에 있어서,

하나 이상의 프로세서;

메모리; 및

상기 메모리에 저장되어 있으며 상기 하나 이상의 프로세서에 의하여 실행되도록 구성되는 하나 이상의 프로그램을 포함하고,

상기 프로그램은,

학습된 모델의 파라미터들을 읽어오는 단계;

상기 파라미터들의 분포에 기초하여 고정 길이의 지수부에 대응하는 가수부의 길이 정보를 포함하는 매핑 테이블을 결정하는 단계; 및

각 파라미터들의 값에 따라 상기 파라미터들의 값들을 고정 길이의 지수부 값 및 가변 길이의 가수부를 포함하는 숫자 포맷으로 변환하는 단계

를 포함하는,

프로세서 장치.

청구항 15

제14항에 있어서,

상기 매핑 테이블은,

상기 파라미터들의 분포에 따라 상기 지수부에 따른 상기 가수부의 길이 정보가 결정되는,

프로세서 장치.

청구항 16

제14항에 있어서,

상기 숫자 포맷으로 변환된 지수부 및 가수부를 메모리에 기록하는 단계

를 더 포함하는,

프로세서 장치.

청구항 17

제14항에 있어서,

상기 파라미터들의 값들을 고정 길이의 지수부 값 및 가변 길이의 가수부를 포함하는 숫자 포맷으로 변환하는 단계는,

상기 가변 길이의 가수부에서 첫 번째 나타나는 1을 제외하여 상기 숫자 포맷으로 변환하는 단계

를 포함하는,

프로세서 장치.

발명의 설명

기 술 분 야

[0001] 실시예는, 파라미터를 가변적으로 표현하는 방법 및 이를 위한 장치에 관한 것이다.

배 경 기 술

[0003] 딥러닝 양자화에서 사용되는 숫자 포맷은 고정 소수점 및 부동 소수점 포맷이 있다. 고정 소수점은 고정된 지수 값으로 모든 비트를 가수부를 표현하는 데에 사용하는 방식이고, 부동 소수점은 고정된 길이의 지수부와 가수부로 그 값을 표현하는 방식으로, 소수점의 위치를 고려하지 않고 표현된 실수와 그 위치를 나타내는 지수를 따로 표현한다.

[0004] 고정 소수점 방식에 대비하여 부동 소수점 방식이 더 많은 범위의 수치를 표현할 수 있고 보다 큰 비트의 범위를 가지게 되나, 연산 속도가 느리기 때문에 별도의 연산 장치를 이용하는 경우가 많다.

발명의 내용

과제의 해결 수단

[0006] 뉴럴 프로세서 장치에 있어서, 변환된 파라미터 값의 고정된 길이의 지수부를 읽어오고 상기 지수부의 값에 매핑된 가수부의 길이 정보를 획득하는 비교기; 상기 변환된 파라미터 값의 가수부를 읽어 오고 상기 가수부의 길이 정보를 이용하여 상기 변환된 파라미터 값의 구조를 변환하는 변환기; 및 상기 지수부의 값에 따른 상기 가

수부의 길이 정보가 매핑되는 매핑 테이블을 포함하는, 뉴럴 프로세서 장치가 제공될 수 있다.

[0007] 상기 매핑 테이블은, 파라미터들의 분포에 따라 상기 지수부의 값에 따른 상기 가수부의 길이 정보가 결정될 수 있다.

[0008] 상기 변환기는, 상기 가수부의 길이 정보에 기초하여 상기 가수부의 값의 자리를 이동시킬 수 있다.

[0009] 상기 변환기는, 상기 변환된 파라미터 값을 표현하기 위해 할당되는 비트 수에 맞추어 0을 추가할 수 있다.

[0010] 상기 변환된 파라미터 값의 표현 방식에 기초하여, 상기 비교기는, 상기 파라미터 값의 고정된 길이의 실수부의 값에 매핑된 상기 실수부의 지수 이동 정보를 획득하고, 상기 변환기는, 상기 실수부의 지수 이동 정보에 따라 상기 실수부의 값을 이동할 수 있다.

[0011] 상기 변환기는, 상기 파라미터 값의 소수부의 값 및 상기 실수부의 지수 이동 정보만큼 이동된 상기 실수부의 값을 합산할 수 있다.

[0012] 상기 비교기 및 상기 변환기는, SRAM으로부터 상기 변환된 파라미터의 지수부 및 가수부를 각각 불러올 수 있다.

[0013] MAC 연산을 위해, 상기 구조가 변환된 상기 파라미터 값 및 상기 지수부의 값이 연산기로 입력될 수 있다.

[0014] 프로세서 장치의 동작 방법에 있어서, 학습된 모델의 파라미터들을 읽어오는 단계; 상기 파라미터들의 분포에 기초하여 고정 길이의 지수부에 대응하는 가수부의 길이 정보를 포함하는 매핑 테이블을 결정하는 단계; 및 각 파라미터들의 값에 따라 상기 파라미터들의 값들을 고정 길이의 지수부 값 및 가변 길이의 가수부를 포함하는 숫자 포맷으로 변환하는 단계를 포함하는, 프로세서 장치의 동작 방법이 제공될 수 있다.

[0015] 상기 매핑 테이블은, 상기 파라미터들의 분포에 따라 상기 지수부에 따른 상기 가수부의 길이 정보가 결정될 수 있다.

[0016] 상기 숫자 포맷으로 변환된 지수부 및 가수부를 메모리에 기록하는 단계를 더 포함할 수 있다.

[0017] 상기 파라미터들의 값들을 고정 길이의 지수부 값 및 가변 길이의 가수부를 포함하는 숫자 포맷으로 변환하는 단계는, 상기 가변 길이의 가수부에서 첫 번째 나타나는 1을 제외하여 상기 숫자 포맷으로 변환하는 단계를 포함할 수 있다.

[0018] 뉴럴 프로세서 장치에 있어서, 하나 이상의 프로세서; 메모리; 및 상기 메모리에 저장되어 있으며 상기 하나 이상의 프로세서에 의하여 실행되도록 구성되는 하나 이상의 프로그램을 포함하고, 상기 프로그램은, 학습된 모델의 파라미터들을 읽어오는 단계; 상기 파라미터들의 분포에 기초하여 고정 길이의 지수부에 대응하는 가수부의 길이 정보를 포함하는 매핑 테이블을 결정하는 단계; 및 각 파라미터들의 값에 따라 상기 파라미터들의 값들을 고정 길이의 지수부 값 및 가변 길이의 가수부를 포함하는 숫자 포맷으로 변환하는 단계를 포함하는, 뉴럴 프로세서 장치가 제공될 수 있다.

도면의 간단한 설명

[0020] 도 1은 실시예에서, 데이터 값을 표현하는 포맷의 형태를 설명하기 위한 도면이다.

도 2는 실시예에서, 뉴럴 프로세서 장치에서 파라미터를 변환하는 방법을 설명하기 위한 흐름도이다.

도 3은 실시예에서, 연산기에서 연산 가능한 형태로 파라미터를 다시 변환하기 위한 뉴럴 프로세서 장치의 구조를 설명하기 위한 도면이다.

도 4는 실시예에서, 연산기에서 연산 가능한 형태로 파라미터를 다시 변환하기 위한 뉴럴 프로세서 장치의 다른 형태의 구조를 설명하기 위한 도면이다.

도 5는 실시예에서, 연산기의 동작을 설명하기 위한 도면이다.

도 6은 실시예에서, 스마트폰 AP로 구현하는 방법을 설명하기 위한 도면이다.

도 7은 실시예에서, 스마트폰 AP로 구현하는 다른 방법을 설명하기 위한 도면이다.

도 8은 일 실시예에 따른 파라미터를 변환하는 프로세서 장치의 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0021] 이하에서, 첨부된 도면을 참조하여 실시예들을 상세하게 설명한다. 그러나, 실시예들에는 다양한 변경이 가해질 수 있어서 특허출원의 권리 범위가 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 실시예들에 대한 모든 변경, 균등물 내지 대체물이 권리 범위에 포함되는 것으로 이해되어야 한다.

[0022] 실시예에서 사용한 용어는 단지 설명을 목적으로 사용된 것으로, 한정하려는 의도로 해석되어서는 안된다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 명세서 상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0023] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0025] 또한, 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일한 참조부호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다. 실시예를 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 실시예의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.

[0026] 또한, 실시 예의 구성 요소를 설명하는 데 있어서, 제1, 제2, A, B, (a), (b) 등의 용어를 사용할 수 있다. 이러한 용어는 그 구성 요소를 다른 구성 요소와 구별하기 위한 것일 뿐, 그 용어에 의해 해당 구성 요소의 본질이나 차례 또는 순서 등이 한정되지 않는다. 어떤 구성 요소가 다른 구성요소에 "연결", "결합" 또는 "접속"된다고 기재된 경우, 그 구성 요소는 그 다른 구성요소에 직접적으로 연결되거나 접속될 수 있지만, 각 구성 요소 사이에 또 다른 구성 요소가 "연결", "결합" 또는 "접속"될 수도 있다고 이해되어야 할 것이다.

[0027] 어느 하나의 실시 예에 포함된 구성요소와, 공통적인 기능을 포함하는 구성요소는, 다른 실시 예에서 동일한 명칭을 사용하여 설명하기로 한다. 반대되는 기재가 없는 이상, 어느 하나의 실시 예에 기재한 설명은 다른 실시 예에도 적용될 수 있으며, 중복되는 범위에서 구체적인 설명은 생략하기로 한다.

[0029] 도 1은 실시예에서, 데이터 값을 표현하는 포맷의 형태를 설명하기 위한 도면이다.

[0030] 이하에서 설명되는 숫자 포맷은 다양한 표현 범위를 가지는 범용적 숫자 포맷이다. 숫자 포맷은 도 1과 같은 형태를 가진다. 부동 소수점의 형태를 가지고 표현하되, 지수부의 길이는 고정되고 가수부의 길이가 가변적인 형태로 표현된다. 지수부 값에 따라 가수부의 길이가 매핑된 길이로 결정될 수 있다.

[0031] 표현하고 싶은 데이터 값의 범위 및 공간에 따라 매핑 테이블이 결정되고, 표현하고자 하는 데이터 값의 범위 및 공간은 응용에 따라 다르게 결정될 수 있다.

[0032] 도 1에 도시된 바와 같은 포맷으로 데이터 값을 표현하는 방법에 대해 설명하고자 한다.

[0033] 지수부는 길이가 고정되기 때문에, 지수부의 경우의 수를 e라고 할 때, $\text{ceil}(\log_2 e)$ 비트가 필요하다. 예컨대, 지수부의 경우의 수가 8개라면 3비트가 필요하게 된다.

[0034] 가수부는 처음 등장하는 1을 빼고 그 뒤에 이어지는 숫자들이 그대로 지수부의 뒤에 이어지게 되며, 0은 특정 지수부 코드로 매핑 테이블에 저장될 수 있다. 지수부의 값에 따라 가수부가 몇 비트인지를 나타내기 위해 별도의 매핑 테이블로 저장되거나 하드웨어로 고정될 수 있다.

[0036] 도 2는 실시예에서, 뉴럴 프로세서 장치에서 파라미터를 변환하는 방법을 설명하기 위한 흐름도이다.

[0037] 이하 실시예에서 각 단계들은 순차적으로 수행될 수도 있으나, 반드시 순차적으로 수행되는 것은 아니다. 예를 들어, 각 단계들의 순서가 변경될 수도 있으며, 적어도 두 단계들이 병렬적으로 수행될 수도 있다.

[0038] 뉴럴 프로세서 장치는 단계(110) 내지 단계(130)을 통해 뉴럴 네트워크의 파라미터를 변환할 수 있다.

[0039] 단계(210)에서 뉴럴 프로세서 장치는, 학습된 모델의 파라미터들을 읽어 온다.

[0040] 해당 파라미터들은 학습된 모델의 적어도 하나의 레이어의 파라미터에 해당할 수 있다.

[0041] 단계(220)에서 뉴럴 프로세서 장치는, 고정 길이의 지수부에 대응하는 가수부의 길이 정보를 포함하는 매핑 테이블을 결정한다.

[0042] 가수부의 길이 정보는, 가수부를 표현하고자 하는 비트 수를 지시할 수 있다.

[0043] 매핑 테이블은 파라미터들의 분포에 기초하여 지수부에 매핑되는 가수부의 길이 정보가 자동으로 결정되거나, 실험적인 분석에 의해 결정될 수 있다. 예컨대, 분포가 높은 파라미터에 대해서는 많은 가수부가 표현되도록 가수부의 길이 정보가 많은 비트들을 할당할 수 있고, 큰 숫자 내에서는 작은 숫자가 생략되도록 할 수 있다.

[0044] 단계(230)에서, 뉴럴 프로세서 장치는, 파라미터의 값들을 고정 길이의 지수부 값 및 가변 길이의 가수부를 포함하는 숫자 포맷으로 변환한다.

[0045] 앞서 도 1을 통해 설명된 숫자 포맷으로 파라미터가 변환될 수 있다. 먼저, 파라미터를 지수부와 가수부로 나누고, 지수부의 경우의 수를 획득할 수 있다. 획득한 경우의 수로 지수부의 값에 대해 ceil($\log_2 e$) 비트로 인코드할 수 있다.

[0046] 이후, 각 지수부 값에 대해 처음 등장하는 1을 제외하고 몇 개의 비트로 가수부를 표현할 지를 지시하는 매핑 테이블을 참조하여 비트 수에 맞게 가수부를 표현하고 인코드된 지수부와 첫 번째 1을 제외한 가수부를 세트로 숫자 포맷을 형성할 수 있다.

[0047] 변환 결과로, 가수부 길이를 나타내는 매핑 테이블과 인코딩된 비트들, 숫자 포맷에 해당하는 값들이 출력될 수 있다.

[0049] 도 3은 실시예에서, 연산기에서 연산 가능한 형태로 파라미터를 다시 변환하기 위한 뉴럴 네트워크 프로세서 장치의 구조를 설명하기 위한 도면이다.

[0050] 뉴럴 네트워크 프로세서 장치(300)는, 변환된 파라미터의 지수부를 읽어오고, 지수부의 값에 따라 매핑 테이블(301)에 매핑된 가수부의 길이 정보를 획득하는 비교기(310)(Comparator) 및 변환된 파라미터의 가수부를 읽어오고 가수부의 값을 획득한 가수부의 길이 정보에 기초하여 가수부의 값을 변환 이전의 값으로 변환하는 변환기(320)(Shifter)를 포함할 수 있다.

[0051] 비교기(310)는 매핑 테이블(301)을 참조하여 지수부의 값에 따른 가수부의 길이 정보를 획득할 수 있다.

[0052] 지수부와 가수부가 따로 저장되어 있는 메모리(예컨대 SRAM 등)를 통해 지수부의 값 및 가수부의 값은 비교기(310) 및 변환기(320) 각각으로 제공될 수 있다.

[0053] 비교기(310)는 지수부의 값과 이에 대응하는 매핑 테이블(301)을 비교하여 가수부가 얼마나 압축되어 있는지 확인할 수 있다. 매핑 테이블(301)은 지수부의 값에 따라 몇 비트로 가수부가 표현되었는지를 나타낸다. 매핑 테이블(301)은 변환 이전의 가수부에서 첫 번째 나타난 1을 제외한 나머지 수치에 대해 몇 비트로 나타낼지를 지시할 수 있다.

[0054] 변환기(320)는 가수부의 값을 지수부의 값에 매핑된 길이 정보만큼 시프트함으로써 변환 이전의 가수부 형태로 변환할 수 있다.

[0055] 이후 연산기(330)로 입력되어 연산 가능한 형태가 될 수 있다. 연산기(330)는 변환기(320)에 의해 변환된 파라미터의 가수부와 지수부의 값을 이용하여 변환 이전의 파라미터 값을 획득하고, 해당 파라미터로 MAC 연산을 실행할 수 있다.

[0057] 도 4는 실시예에서, 연산기에서 연산 가능한 형태로 파라미터를 다시 변환하기 위한 뉴럴 네트워크 프로세서 장치의 다른 형태의 구조를 설명하기 위한 도면이다.

[0058] 도 4에 도시된 구조는 고정 소수점 형태로 파라미터가 표현된 경우, 변환된 파라미터를 다시 변환 이전의 형태

를 가지는 파라미터로 변환하기 위한 것이다. 도 4의 예시는, 지수부가 4비트인 경우로 설명하고자 한다.

[0059] 뉴럴 네트워크 프로세서 장치(400)는, 고정 소수점으로 변환된 파라미터의 실수부를 읽어오고, 실수부의 값에 따라 매핑 테이블(401)에 매핑된 실수부의 지수 이동 정보를 획득하는 비교기(410) 및 고정 소수점으로 변환된 파라미터의 소수부를 읽어오고 소수부의 값을 지수 이동 정보에 따라 지수가 이동된 실수부의 값과 합산하여 변환 이전의 파라미터로 변환시키는 변환기(420)를 포함할 수 있다.

[0060] 도 4에 따르면, 변환기(420)로부터 고정 소수점 형태로 변환된 파라미터에 대한 변환 이전의 파라미터 값을 획득할 수 있다.

[0061] 매핑 테이블(401)은 실수부의 값에 따라 해당 실수부에 대한 지수 이동 정보를 매핑할 수 있다. 지수 이동 정보는 실수부의 값이 감소함에 따라 지수 이동 정보가 1자리씩 줄어드는 형태로 제공될 수 있다. 예컨대, 8비트의 파라미터에 대해서는 111 내지 000까지의 실수부 값이 감소함에 따라 1xxxxxxx, 1xxxxxx, 1xxxxx, ... , 1의 형태로 줄어드는 매핑 테이블이 제공될 수 있다.

[0062] 혹은 매핑 테이블(401)의 매핑된 값은 실수부의 값에 따라, 실수부의 오른쪽으로 이동해야 하는 자릿수 나타내는 것으로 표현될 수도 있다. 예컨대, 실수부의 값이 010인 경우, 소수부의 값을 실수부의 1자리 이후 표현될 수 있다.

[0063] 변환기(420)는 비교기(410)로부터 획득한 실수부의 값 및 실수부의 지수 이동 정보에 기초하여 실수부의 지수를 이동시키고, 소수부의 값을 그대로 합산함으로써 변환 이전의 16비트의 파라미터를 획득할 수 있다.

[0064] SRAM와 같은 메모리에 실수부 및 소수부의 정보가 각각 저장될 수 있고, 각각 저장된 정보가 비교기(410) 및 변환기(420)로 각각 입력될 수 있다.

[0065] 연산 가능한 형태의 변환 이전의 파라미터는 연산기(430)로 입력되어 MAC 연산이 이루어질 수 있다. 연산기(430)는 변환기(320)에 의해 변환 이전의 파라미터를 이용하여 MAC 연산을 실행할 수 있다.

[0067] 도 5는 실시예에서, 연산기의 동작을 설명하기 위한 도면이다.

[0068] 연산기에서 16비트의 파라미터를 이용하여 MAC 연산을 실행하는 방법에 대해서 설명하도록 한다.

[0069] 변환 이전의 파라미터로 가중치(Weight)가 설정될 수 있고, 입력되는 활성화(Activation)에 대해서 MAC 연산을 수행할 수 있다.

[0070] 가중치와 활성화의 데이터 포맷을 FP16(half-precision floating point)의 이진 형식이라고 가정하고, 가중치는 1.25: 0011_1101_0000_0000, 활성화는 2.125: 0100_0000_1000_0000 가정할 수 있다.

[0071] 가중치를 분석하면, 부호: 0, 지수부: 01111, 가수부: 0100000000이고, 이 값을 매핑 테이블에 지수부는 01111, 변환된 가수부는 0000000101로 매핑 테이블에 표현될 수 있다. 해당 매핑 테이블로부터 가수부를 01로 변환하여 메모리에 저장할 수 있게 된다.

[0072] 연산기에 입력하기 위해서는 변환된 파라미터로부터 변환 이전의 파라미터를 획득한다. 변환된 파라미터로부터 지수부 01111를 획득하여 매핑 테이블에 따라 가수부의 이동 정보가 8칸임을 확인할 수 있다.

[0073] 가중치를 변환 이전의 값으로 변환한 후, 도 5와 같이 연산을 수행할 수 있다. 가중치 및 활성화 각각을 부호, 지수부, 가수부로 나눈 후, 각각에 대해서 연산을 수행한다.

[0074] 부호는 XOR 연산을 수행하고, 지수부는 더하기를 수행하고, 가수부는 곱하기를 수행한다. 이후 지수부의 연산 결과와 가수부의 연산 결과를 정규화(Normalization) 및 반올림(Rounding)을 수행하여 출력(Output)을 획득할 수 있다.

[0075] 획득한 출력은, 다음 레이어의 활성화가 될 수 있고, 또는 최종 출력 값이 될 수 있다.

[0077] 도 6은 실시예에서, 스마트폰 AP로 구현하는 방법을 설명하기 위한 도면이다.

[0078] 도 6은 스마트폰 AP의 내부 구성을 개략적으로 나타낸 것이다. 스마트폰 AP는 NPU(Neural Processing Unit)를 포함할 수 있다. NPU는 앞서 설명된 숫자 포맷으로 저장된 딥러닝 파라미터를 연산하도록 구비될 수 있다.

[0079] 고정 지수부 및 가변 가수부로 표현되는 숫자 포맷을 연산하기 위한 NPU 연산기(Arithmetic Unit)를 구성함으로써, 기존의 고정 소수점 포맷을 포함하여 도 1 내지 도 5의 설명을 통해 정의된 포맷들에 대해 곱셈 연산들을 처리할 수 있다.

[0080] 딥러닝 파라미터들의 분포가 고정 지수부 및 가변 가수부로 표현되는 숫자 포맷의 효율적인 크기로 SRAM 등의 온칩메모리(On-chip memory)에 저장될 수 있고, 탄력적으로 정확도와 연산 효율, 저장 효율, 공간 효율이 조절될 수 있다.

[0082] 도 7은 실시예에서, 스마트폰 AP로 구현하는 다른 방법을 설명하기 위한 도면이다.

[0083] 도 6과 달리, 도 7의 스마트폰 AP는 스마트폰의 AP와 같은 곳에서 인코딩이 활용되고, 디코더(701)가 내장되는 구성을 도시하고 있다. 이런 방식으로 프로세서를 구성하게 되면, 연산기의 구성을 변형하지 않고 숫자 포맷만 실시예의 방식대로 저장할 수 있다.

[0084] 같은 고정 소수점 표현 공간을 저장하더라도, 딥러닝 파라미터의 분포 특성 상, 작은 크기로 메모리에 저장될 수 있다.

[0085] 스마트폰의 NPU가 신경망을 연산할 때에 외부 메모리에서 파라미터 정보를 읽어오는 부분에 많은 전력이 소모되므로, 변환된 파라미터를 읽어오게 되면, 적은 비트수로 표현되는 실시예의 파라미터를 통해 파라미터에 대한 통신량 및 소모 전력을 줄일 수 있다.

[0086] 만약, 파라미터를 읽어오는 데에 신경망 처리 속도가 저하되는 경우, 신경망 처리 속도를 높일 수 있다.

[0088] 도 8은 일 실시예에 따른 파라미터를 변환하는 프로세서 장치의 블록도이다.

[0089] 도 6를 참조하면, 일 실시예에 따른 프로세서 장치(800)는 통신 인터페이스(810), 프로세서(830), 및 메모리(850)를 포함할 수 있다. 통신 인터페이스(810), 프로세서(830) 및 메모리(850)는 통신 버스(805)를 통해 서로 통신할 수 있다.

[0090] 통신 인터페이스(810)는 파라미터를 수신한다.

[0091] 프로세서(830)는 통신 인터페이스(810)를 통해 수신된 파라미터를 미리 정해진 숫자 포맷으로 변환할 수 있다. 프로세서(830)는 획득한 파라미터를 매핑 테이블을 참조하여 고정된 지수부 및 가변적 가수부를 포함하는 숫자 포맷으로 변환할 수 있다.

[0092] 메모리(850)는 상술한 프로세서(830)의 처리 과정에서 생성되는 다양한 정보들을 저장할 수 있다. 이 밖에도, 메모리(850)는 각종 데이터와 프로그램 등을 저장할 수 있다. 메모리(850)는 휘발성 메모리 또는 비휘발성 메모리를 포함할 수 있다. 메모리(850)는 하드 디스크 등과 같은 대용량 저장 매체를 구비하여 각종 데이터를 저장할 수 있다.

[0093] 또한, 프로세서(830)는 도 1 내지 도 7을 통해 전술한 적어도 하나의 방법 또는 적어도 하나의 방법에 대응되는 알고리즘을 수행할 수 있다. 프로세서(530)는 목적하는 동작들(desired operations)을 실행시키기 위한 물리적인 구조를 갖는 회로를 가지는 하드웨어로 구현된 데이터 처리 장치일 수 있다. 예를 들어, 목적하는 동작들은 프로그램에 포함된 코드(code) 또는 인스트럭션들(instructions)을 포함할 수 있다. 프로세서(830)는 예를 들어, CPU(Central Processing Unit), GPU(Graphics Processing Unit), 또는 NPU(Neural network Processing Unit)으로 구성될 수 있다. 예를 들어, 하드웨어로 구현된 예측 장치(500)는 마이크로프로세서(microprocessor), 중앙 처리 장치(central processing unit), 프로세서 코어(processor core), 멀티-코어 프로세서(multi-core processor), 멀티프로세서(multiprocessor), ASIC(Application-Specific Integrated Circuit), FPGA(Field Programmable Gate Array)를 포함할 수 있다.

[0094] 프로세서(830)는 프로그램을 실행하고, 프로세서 장치(800)를 제어할 수 있다. 프로세서(830)에 의하여 실행되는 프로그램 코드는 메모리(850)에 저장될 수 있다.

[0096] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판

독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

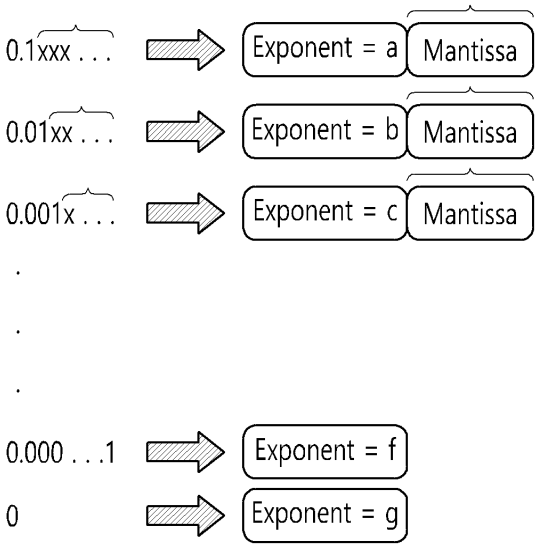
[0098] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0100] 이상과 같이 실시예들이 비록 한정된 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기를 기초로 다양한 기술적 수정 및 변형을 적용할 수 있다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

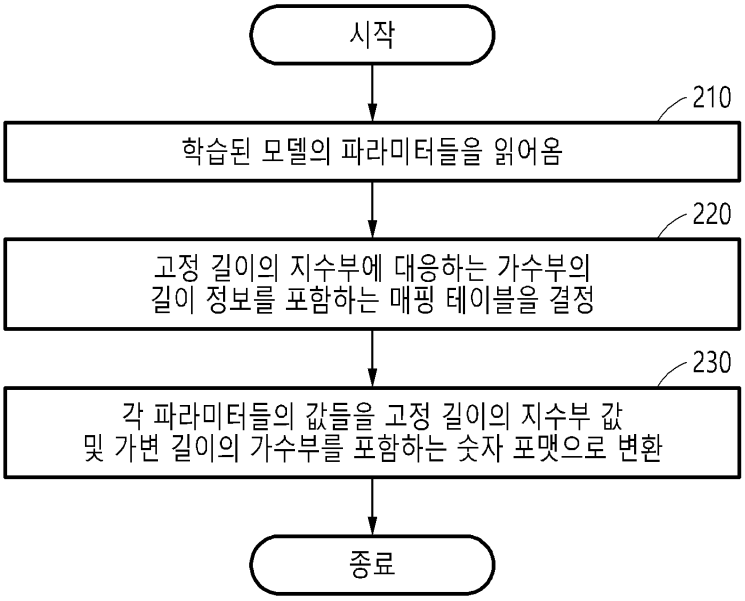
[0101] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 청구범위의 범위에 속한다.

도면

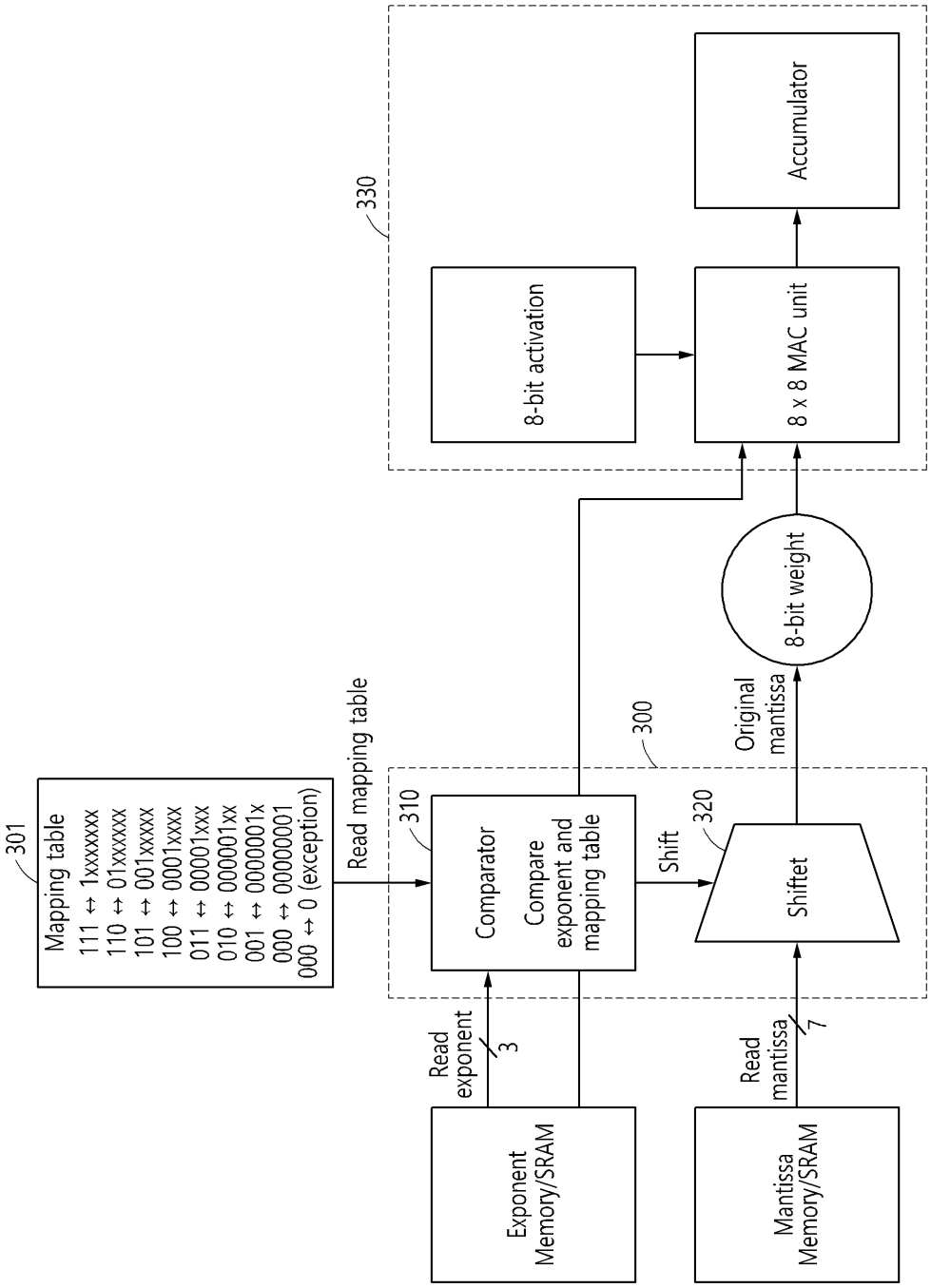
도면1



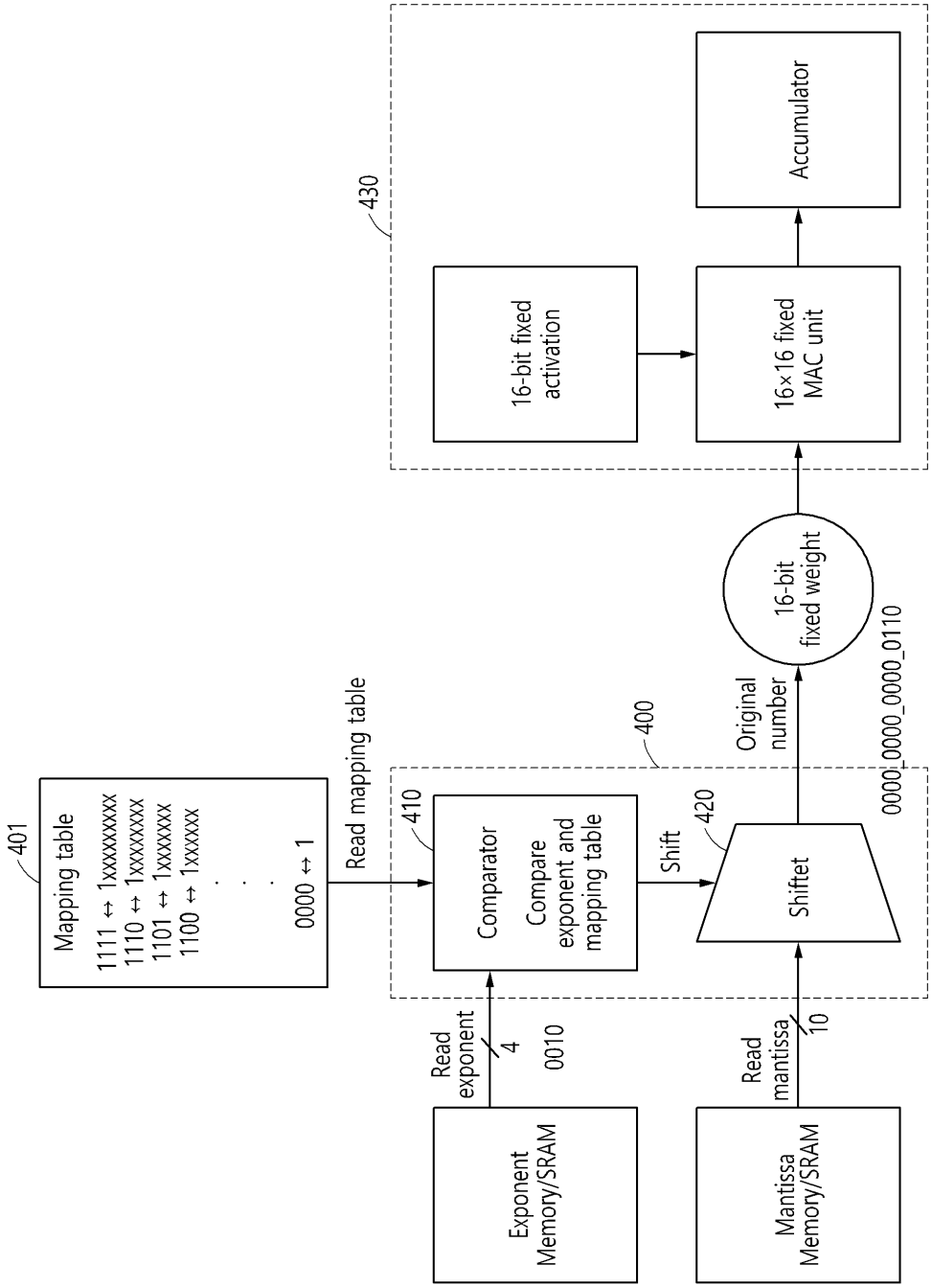
도면2


시작
210
학습된 모델의 파라미터들을 읽어옴
220
고정 길이의 지수부에 대응하는 가수부의
길이 정보를 포함하는 매핑 테이블을 결정
230
각 파라미터들의 값들을 고정 길이의 지수부 값
및 가변 길이의 가수부를 포함하는 숫자 포맷으로 변환
종료

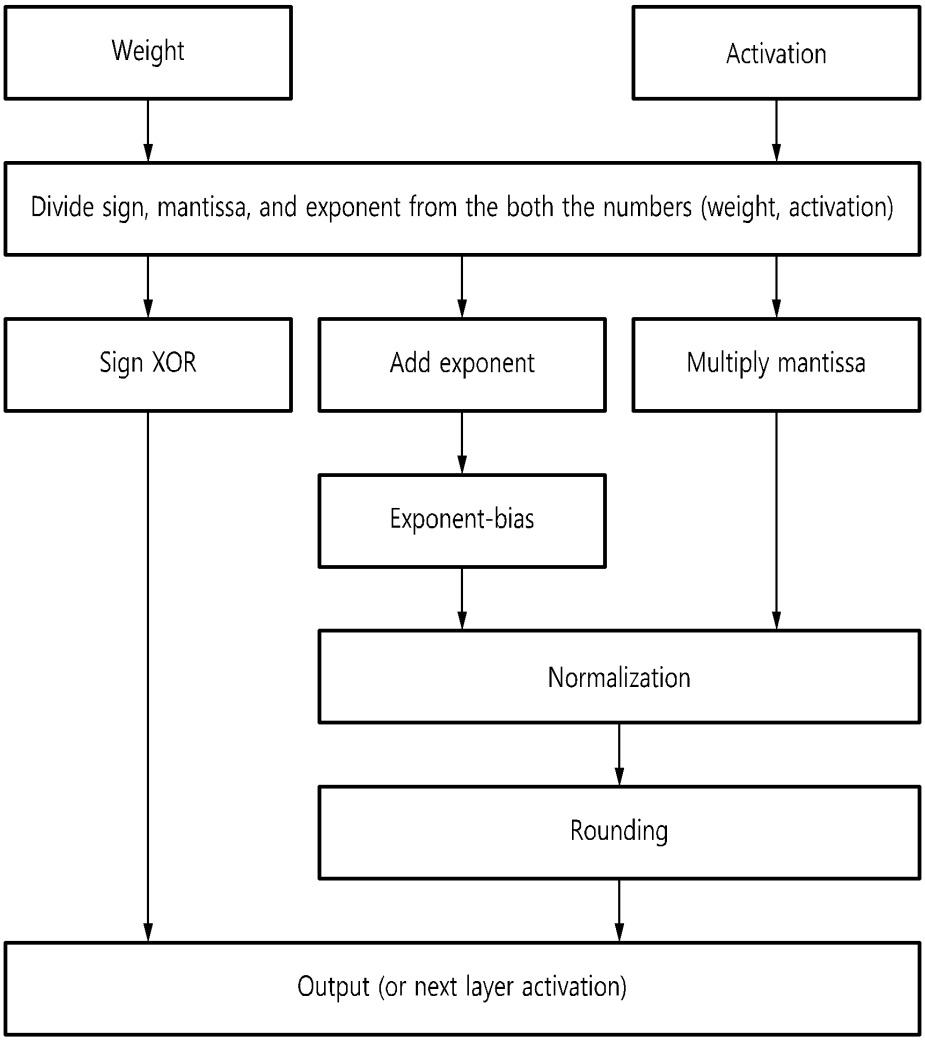
도면3


330
Accumulator
8-bit activation
8 x 8 MAC unit
8-bit weight
300
Original mantissa
301
Mapping table
111 ↔ 1xxxxxxx
110 ↔ 01xxxxxx
101 ↔ 001xxxxx
100 ↔ 0001xxxx
011 ↔ 00001xxx
010 ↔ 000001xx
001 ↔ 0000001x
000 ↔ 00000001
000 ↔ 0 (exception)
Read mapping table
310
Comparator
Compare exponent and mapping table
Shift
320
Shiftet
Read exponent
3
Read mantissa
7
Exponent Memory/SRAM
Mantissa Memory/SRAM

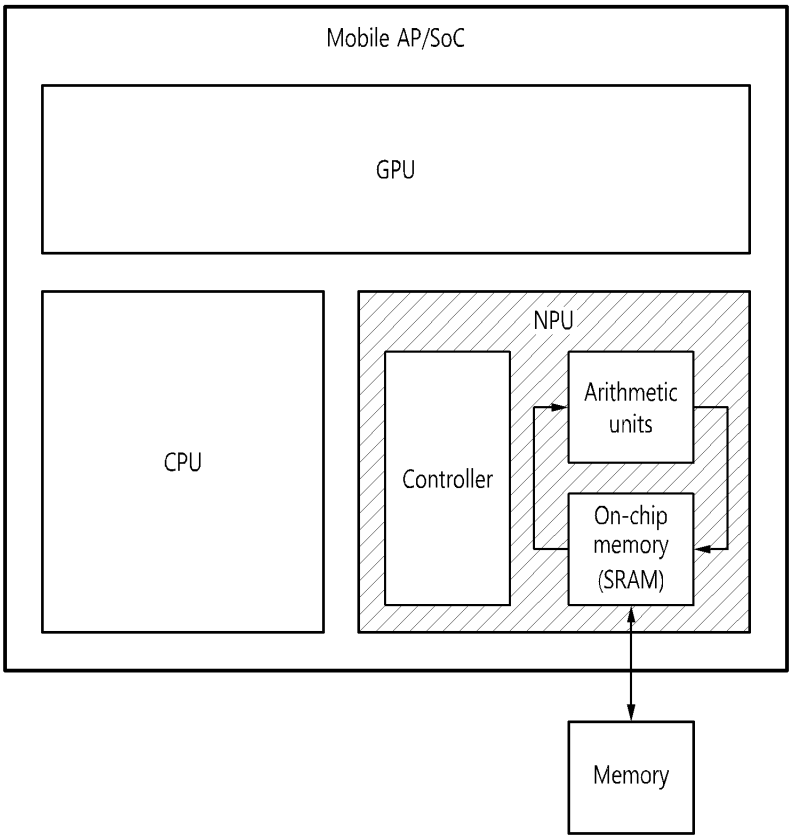
도면4


401
Mapping table
1111 ↔ 1xxxxxxxxx
1110 ↔ 1xxxxxxxx
1101 ↔ 1xxxxxxx
1100 ↔ 1xxxxxx
0000 ↔ 1
Read mapping table
400
410
Comparator
Compare exponent and mapping table
Shift
420
Shiftet
Original number
0000_0000_0000_0110
16-bit fixed weight
430
16-bit fixed activation
16×16 fixed MAC unit
Accumulator
Read exponent
4
0010
Exponent Memory/SRAM
Read mantissa
10
Mantissa Memory/SRAM

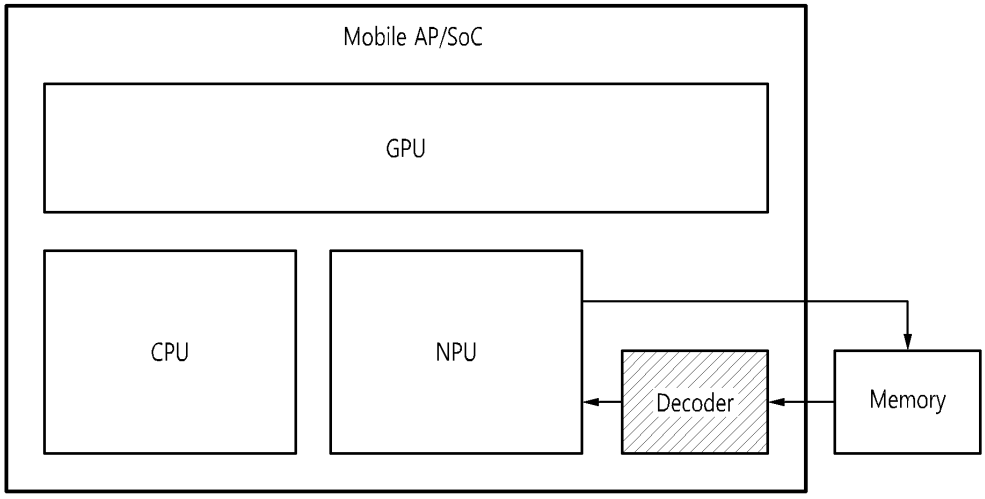
도면5


Weight
Activation
Divide sign, mantissa, and exponent from the both the numbers (weight, activation)
Sign XOR
Add exponent
Multiply mantissa
Exponent-bias
Normalization
Rounding
Output (or next layer activation)

도면6


Mobile AP/SoC
GPU
CPU
NPU
Controller
Arithmetic units
On-chip memory (SRAM)
Memory

도면7


Mobile AP/SoC
GPU
CPU
NPU
Decoder
Memory

도면8

