

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0157799

(43) 공개일자 2022년11월29일

(51) 국제특허분류(Int. Cl.)

G06F 7/544 (2017.01) G06F 7/53 (2006.01)

G06T 3/40 (2006.01) G06T 7/11 (2017.01)

(52) CPC특허분류

G06F 7/5443 (2013.01)

G06F 7/5312 (2013.01)

(21) 출원번호 10-2021-0065778

(22) 출원일자 2021년05월21일

심사청구일자 2021년05월21일

(71) 출원인

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

정성욱

서울특별시 서대문구 연세로 50, 제3공학관 C513 (신촌동, 연세대학교)

안홍근

서울특별시 서대문구 연세로 50, 제3공학관 C421 (신촌동, 연세대학교)

(뒷면에 계속)

(74) 대리인

특허법인우인

전체 청구항 수 : 총 13 항

(54) 발명의 명칭 인공지능 알고리즘의 효율적인 하드웨어 구현을 위한 수용 영역의 분할 연산 방법

(57) 요약

본 실시예들은 이미지를 입력하는 학습 네트워크 모델을 처리할 때 수용 영역을 복수의 지역 영역으로 나누어 연산을 진행하며, 지역 영역의 불연속으로 인한 각 데이터의 끝부분에서 불완전한 데이터를 임시 저장하고, 임시 저장된 불완전한 결과 데이터를 반영하여 완전한 데이터에 대해서 다음 곱셈 누적 연산을 수행할 수 있는 수용 영역의 분할 연산 방법 및 지역 영역 기반의 곱셈 누적 연산 장치를 제공한다.

대표도 - 도7



(52) CPC특허분류

G06N 3/063 (2013.01)

G06T 3/4053 (2013.01)

G06T 7/11 (2017.01)

(72) 발명자

주성환

서울특별시 서대문구 연세로 50, 제3공학관 C421(
신촌동, 연세대학교)

이수민

서울특별시 서대문구 연세로 50, 제3공학관 C421(
신촌동, 연세대학교)

정영석

서울특별시 서대문구 연세로 50, 제3공학관 C421(
신촌동, 연세대학교)

명세서

청구범위

청구항 1

수용 영역의 분할 연산 방법에 있어서,
입력 데이터가 저장된 메모리에 할당된 수용 영역을 복수의 지역 영역으로 분할하는 단계; 및
상기 복수의 지역 영역에 대해서 연산을 수행하는 단계를 포함하는 수용 영역의 분할 연산 방법.

청구항 2

제1항에 있어서,
상기 연산을 수행하는 단계는,
상기 복수의 지역 영역 중에서 이전 지역 영역에 대해서 연산을 수행한 불완전한 결과 데이터를 임시 저장하는 단계;
상기 복수의 지역 영역 중에서 다음 지역 영역이 시작되는 순간 상기 임시 저장된 불완전한 결과 데이터를 불러 오는 단계; 및
상기 임시 저장된 불완전한 결과 데이터를 반영하여 완전한 데이터에 대해서 곱셈 누적 연산을 수행하는 단계를 포함하는 수용 영역의 분할 연산 방법.

청구항 3

제1항에 있어서,
상기 연산을 수행하는 단계는 이미지의 픽셀 데이터에 대해서 커널 단위로 곱셈 누적 연산을 수행하는 것을 특징으로 하는 수용 영역의 분할 연산 방법.

청구항 4

제1항에 있어서,
상기 연산을 수행하는 단계는 이미지의 픽셀 데이터에 대해서 초해상도 변환을 수행하는 것을 특징으로 하는 수용 영역의 분할 연산 방법.

청구항 5

제1항에 있어서,
상기 연산을 수행하는 단계는 학습 네트워크 모델을 통해 입력 처리 단위에 해당하는 상기 수용 영역으로부터 출력 처리 단위에 해당하는 투사 영역으로 연산을 수행하며, 상기 지역 영역 단위를 기준으로 메모리를 접근하는 것을 특징으로 하는 수용 영역의 분할 연산 방법.

청구항 6

지역 영역 기반의 곱셈 누적 연산 장치에 있어서,
제1 데이터와 제2 데이터를 곱셈하여 곱셈 결과를 출력하는 곱셈부;
상기 곱셈 결과를 제1 경로와 제2 경로로 분할하는 분할부;
상기 제1 경로에 따른 제1 곱셈 결과를 누적하고 저장하는 제1 덧셈부;
상기 제2 경로에 따른 제2 곱셈 결과를 누적하고 저장하는 제2 덧셈부; 및
상기 분할부, 상기 제1 덧셈부, 상기 제2 덧셈부에 제어 신호를 송신하여 동작을 제어하는 제어부를 포함하는

지역 영역 기반의 곱셈 누적 연산 장치.

청구항 7

제6항에 있어서,

상기 누적된 제2 곱셈 결과를 가공하는 후처리부를 포함하는 것을 특징으로 하는 지역 영역 기반의 곱셈 누적 연산 장치.

청구항 8

제6항에 있어서,

상기 제2 덧셈부는 다음 지역 영역 사이클에서 상기 제1 덧셈부의 상기 누적된 제1 곱셈 결과를 고려하는 것을 특징으로 하는 지역 영역 기반의 곱셈 누적 연산 장치.

청구항 9

제6항에 있어서,

상기 분할부, 상기 제1 덧셈부, 상기 제2 덧셈부를 통해 입력 데이터가 저장된 메모리에 할당된 수용 영역을 상기 지역 영역으로 분할하여 상기 제1 데이터와 상기 제2 데이터에 대한 연산을 수행하는 것을 특징으로 하는 지역 영역 기반의 곱셈 누적 연산 장치.

청구항 10

제6항에 있어서,

상기 제어부는 상기 분할부로 제1 선택 신호를 송신하고,

상기 제1 선택 신호에 따라 상기 분할부는 상기 제1 경로 또는 상기 제2 경로를 활성화시키는 것을 특징으로 하는 지역 영역 기반의 곱셈 누적 연산 장치.

청구항 11

제6항에 있어서,

상기 제어부는 상기 제1 덧셈부 및 상기 제2 덧셈부로 클럭 신호를 송신하고, 상기 클럭 신호에 따라 상기 제1 덧셈부 및 상기 제2 덧셈부가 동기화되는 것을 특징으로 하는 지역 영역 기반의 곱셈 누적 연산 장치.

청구항 12

제6항에 있어서,

상기 제어부는 상기 제1 덧셈부로 리셋 신호를 송신하고,

상기 리셋 신호에 따라 상기 제1 덧셈부는 임시 저장된 데이터를 초기화하는 것을 특징으로 하는 지역 영역 기반의 곱셈 누적 연산 장치.

청구항 13

제6항에 있어서,

상기 제어부는 상기 제2 덧셈부로 제2 선택 신호를 송신하고,

상기 제2 선택 신호에 따라 상기 제2 덧셈부는 상기 제1 덧셈부에 임시 저장된 데이터를 수신하는 것을 특징으로 하는 지역 영역 기반의 곱셈 누적 연산 장치.

발명의 설명

기술 분야

본 발명이 속하는 기술 분야는 수용 영역의 분할 연산 방법 및 지역 영역 기반의 곱셈 누적 연산 장치에 관한 것이다. 본 연구는 삼성전자 미래기술육성센터의 지원을 받아 수행된 실시간 초해상도 복원작업에 적용 가능한

[0001]

저전력 신경망 단일 칩 기술개발과 관련된다(No. SRFC-IT1802-06).

배경 기술

- [0002] 이 부분에 기술된 내용은 단순히 본 실시예에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.
- [0003] 인공지능 알고리즘 중에서 CNN(Convolutional Neural Network)은 이미지 처리에 널리 쓰이고 있으며, 해당 알고리즘의 복잡성은 하드웨어 구현에 큰 영향을 끼친다.
- [0004] 시스템에서 한 번에 처리하는 데이터(수용 영역)를 크게 설정하는 경우, 시스템 측면에서 동일한 연산을 중복하여 반복하는 것이 줄게 되어 시스템 성능이 증가하게 된다. 하지만 기존의 연산 구조를 그대로 사용하면 수용 영역 단위로 메모리 접근을 진행하여야 하기 때문에 메모리 बैं드폭 문제에 직면하게 되는 문제가 있다.

선행기술문헌

특허문헌

- [0005] (특허문헌 0001) KR 10-2017-0099848 (2017.09.01)
- (특허문헌 0002) KR 10-2020-0110165 (2020.09.23)
- (특허문헌 0003) KR 10-2020-0144276 (2020.12.29)

발명의 내용

해결하려는 과제

- [0006] 본 발명의 실시예들은 불완전 데이터의 곱셈 누산(Multiply-Accumulate, MAC) 결과물을 MAC 모듈 내부에 임시 저장하고, 다음 지역 영역(local field) 연산에 이를 다시 MAC 내부 누산기(ACC)의 초기값으로 가져오는 방식을 통하여 수용 영역(receptive field)을 분할하여 연산하는데 주된 목적이 있다.
- [0007] 본 발명의 명시되지 않은 또 다른 목적들은 하기의 상세한 설명 및 그 효과로부터 용이하게 추론할 수 있는 범위 내에서 추가적으로 고려될 수 있다.

과제의 해결 수단

- [0008] 본 실시예의 일 측면에 의하면, 수용 영역의 분할 연산 방법에 있어서, 입력 데이터가 저장된 메모리에 할당된 수용 영역을 복수의 지역 영역으로 분할하는 단계; 및 상기 복수의 지역 영역에 대해서 연산을 수행하는 단계를 포함하는 수용 영역의 분할 연산 방법을 제공한다.
- [0009] 상기 연산을 수행하는 단계는, 상기 복수의 지역 영역 중에서 이전 지역 영역에 대해서 연산을 수행한 불완전한 결과 데이터를 임시 저장하는 단계; 상기 복수의 지역 영역 중에서 다음 지역 영역이 시작되는 순간 상기 임시 저장된 불완전한 결과 데이터를 불러오는 단계; 및 상기 임시 저장된 불완전한 결과 데이터를 반영하여 완전한 데이터에 대해서 곱셈 누적 연산을 수행하는 단계를 포함할 수 있다.
- [0010] 상기 연산을 수행하는 단계는 이미지의 픽셀 데이터에 대해서 커널 단위로 곱셈 누적 연산을 수행할 수 있다.
- [0011] 상기 연산을 수행하는 단계는 이미지의 픽셀 데이터에 대해서 초해상도 변환을 수행할 수 있다.
- [0012] 상기 연산을 수행하는 단계는 학습 네트워크 모델을 통해 입력 처리 단위에 해당하는 상기 수용 영역으로부터 출력 처리 단위에 해당하는 투사 영역으로 연산을 수행하며, 상기 지역 영역 단위를 기준으로 메모리를 접근할 수 있다.
- [0013] 본 실시예의 다른 측면에 의하면, 지역 영역 기반의 곱셈 누적 연산 장치에 있어서, 제1 데이터와 제2 데이터를 곱셈하여 곱셈 결과를 출력하는 곱셈부; 상기 곱셈 결과를 제1 경로와 제2 경로로 분할하는 분할부; 상기 제1 경로에 따른 제1 곱셈 결과를 누적하고 저장하는 제1 덧셈부; 상기 제2 경로에 따른 제2 곱셈 결과를 누적하고 저장하는 제2 덧셈부; 및 상기 분할부, 상기 제1 덧셈부, 상기 제2 덧셈부에 제어 신호를 송신하여 동작을 제어하는 제어부를 포함하는 지역 영역 기반의 곱셈 누적 연산 장치를 제공한다.

- [0014] 상기 지역 영역 기반의 곱셈 누적 연산 장치는 상기 누적된 제2 곱셈 결과를 가공하는 후처리부를 포함할 수 있다.
- [0015] 상기 제2 덧셈부는 다음 지역 영역 사이클에서 상기 제1 덧셈부의 상기 누적된 제1 곱셈 결과를 고려할 수 있다.
- [0016] 상기 분할부, 상기 제1 덧셈부, 상기 제2 덧셈부를 통해 입력 데이터가 저장된 메모리에 할당된 수용 영역을 상기 지역 영역으로 분할하여 상기 제1 데이터와 상기 제2 데이터에 대한 연산을 수행할 수 있다.
- [0017] 상기 제어부는 상기 분할부로 제1 선택 신호를 송신하고, 상기 제1 선택 신호에 따라 상기 분할부는 상기 제1 경로 또는 상기 제2 경로를 활성화시킬 수 있다.
- [0018] 상기 제어부는 상기 제1 덧셈부 및 상기 제2 덧셈부로 클록 신호를 송신하고, 상기 클록 신호에 따라 상기 제1 덧셈부 및 상기 제2 덧셈부가 동기화될 수 있다.
- [0019] 상기 제어부는 상기 제1 덧셈부로 리셋 신호를 송신하고, 상기 리셋 신호에 따라 상기 제1 덧셈부는 임시 저장된 데이터를 초기화할 수 있다.
- [0020] 상기 제어부는 상기 제2 덧셈부로 제2 선택 신호를 송신하고, 상기 제2 선택 신호에 따라 상기 제2 덧셈부는 상기 제1 덧셈부에 임시 저장된 데이터를 수신할 수 있다.

발명의 효과

- [0021] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 이미지를 입력하는 학습 네트워크 모델을 처리할 때 수용 영역을 복수의 지역 영역으로 나누어 연산을 진행하며, 지역 영역의 불연속으로 인한 각 데이터의 끝부분에서 불완전한 데이터를 임시 저장하고, 임시 저장된 불완전한 결과 데이터를 반영하여 완전한 데이터에 대해서 다음 곱셈 누적 연산을 수행할 수 있는 효과가 있다.
- [0022] 여기에서 명시적으로 언급되지 않은 효과라 하더라도, 본 발명의 기술적 특징에 의해 기대되는 이하의 명세서에서 기재된 효과 및 그 잠정적인 효과는 본 발명의 명세서에 기재된 것과 같이 취급된다.

도면의 간단한 설명

- [0023] 도 1은 학습 네트워크 모델 기반의 초해상도 기법을 예시한 도면이다.
- 도 2 및 도 3은 이미지에 대한 수용 영역과 투사 영역을 예시한 도면이다.
- 도 4는 이미지의 영역에 따른 메모리 사이즈를 예시한 도면이다.
- 도 5는 기존의 곱셈 누적 연산 장치를 예시한 도면이다.
- 도 6은 메모리에 할당되는 수용 영역, 지역 영역, 투사 영역을 예시한 도면이다.
- 도 7 및 도 8은 본 발명의 일 실시예에 따른 지역 영역 기반의 곱셈 누적 연산 장치를 예시한 도면이다.
- 도 9는 본 발명의 일 실시예에 따른 지역 영역 기반의 곱셈 누적 연산 장치가 데이터를 처리하는 타이밍을 예시한 도면이다.
- 도 10 및 도 11은 본 발명의 다른 실시예에 따른 수용 영역의 분할 연산 방법을 예시한 흐름도이다.
- 도 12는 본 발명의 실시예들에 따른 시뮬레이션 결과를 예시한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0024] 이하, 본 발명을 설명함에 있어서 관련된 공지기능에 대하여 이 분야의 기술자에게 자명한 사항으로서 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하고, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다.
- [0025] 도 1은 학습 네트워크 모델 기반의 초해상도 기법을 예시한 도면이다.
- [0026] 많은 데이터를 이용해 학습하는 방식인 딥러닝(Deep Learning)은 기존 다른 학습 알고리즘보다 훨씬 더 뛰어난 성능을 보여준다. 그 중 CNN(Convolutional Neural Network)은 이미지 관련 트레이닝에서 우수한 성능을 보여주고 널리 활용되고 있다. 이러한 딥러닝은 큰 계산 복잡도 때문에 학습에 적합한 효율적인 하드웨어를 구현하기

곤란한 문제가 있다.

- [0027] CNN 기반의 초해상도 변환 네트워크를 예로 들면, 1 깊이(depth)의 원본 이미지는 초해상도 변환 네트워크를 통해 다수의 레이어를 통과하며, 최종 레이어를 지나고 나면 그 결과는 4 깊이로 확장된 이미지가 생성된다. 4 깊이의 이미지는 리매핑(remapping)을 통해 1 깊이의 이미지로 변환되며, 이는 원본 이미지의 가로, 세로 각각 2배씩 확장된 이미지가 된다. 도 1을 참조하면, FHD(1920x1080) 이미지를 UHD(3840x2160)으로 확장하는 것을 예시로 도시하고 있다.
- [0028] 도 2 및 도 3은 이미지에 대한 수용 영역과 투사 영역을 예시한 도면이다.
- [0029] 고해상도 이미지는 많은 수의 픽셀을 가지고, 이를 한번에 네트워크를 통과시킬 수 없다. 예컨대, FHD=1920x1080 pixels 이를 한 번에 처리하려고 하면, 알고리즘을 수행하기 위해 수 천 개 이상의 PE와 수 백 MB 이상의 bandwidth를 지원하는 메모리가 필요하게 된다. 따라서, 적당한 개수의 처리 단위(Processing Element, PE)를 갖는 처리 장치를 통해 고해상도의 이미지를 순차적/반복적으로 처리해야 한다. 이 때, 한 번 처리하는 입력/출력의 단위는 수용 영역(receptive field)/투사 영역(projective field)이라고 한다.
- [0030] 수용 영역이 작은 경우, 한 번에 처리하는 연산이 적지만 그만큼 많이 반복 수행해야 한다.
- [0031] 수용 영역은 투사 영역과 알고리즘에 따라 결정되고, 반복 횟수는 투사 영역에 따라 결정된다.
- [0032] 도 4는 이미지의 영역에 따른 메모리 사이즈를 예시한 도면이고, 도 5는 기존의 곱셈 누적 연산 장치를 예시한 도면이고, 도 6은 메모리에 할당되는 수용 영역, 지역 영역, 투사 영역을 예시한 도면이다.
- [0033] 영역(field)을 크게 키우면 중복 데이터(overlap data)의 비중이 감소하여 시스템 효율이 증가하지만 영역이 커짐에 따라 메모리 밴드폭 측면에서 악화되는 문제가 있다.
- [0034] 기존의 곱셈 누적 연산 장치를 통해서도 수용 영역과 투사 영역 두 가지 영역을 고려하므로, 기존 연산 구조만 가지고는 영역을 키우는데 있어 메모리 밴드폭 차원에서 한계에 도달하게 된다.
- [0035] 본 명세서에서 제안하는 수용 영역의 분할 연산 방법 및 지역 영역 기반의 곱셈 누적 연산 장치를 통해 수용 영역을 키워 시스템 성능을 높이고, 수용 영역을 나누어 계산할 수 있으므로 메모리 밴드폭을 줄일 수 있다.
- [0036] 본 실시예에 따른 지역 영역 기반의 곱셈 누적 연산 장치는 수용 영역을 여러 개의 지역 영역(local field)으로 나누어 연산을 진행한다. 지역 영역 기반의 곱셈 누적 연산 장치는 수용 영역 데이터를 순차적으로(Serial) 처리 가능하여, 메모리 밴드폭 문제를 완화하고 시스템 성능을 향상시킬 수 있다.
- [0037] 메모리 특성 상 동시에 서로 다른 워드라인(WL)에 있는 데이터에 접근할 수 없으므로, 각 지역 영역(local field)는 도 6과 같이 서로 독립적인 데이터를 가지게 된다.
- [0038] 커널을 이용한 컨볼루션 연산을 진행하게 되면 지역 영역의 불연속으로 인해 각 데이터의 끝 부분에서 온전하지 않은 데이터가 MAC 모듈로 들어가는 경우가 발생하게 된다. 예컨대, 커널(Kernel)은 3x3인데 비해, 데이터는 일부분만 (2x3, 또는 1x3) 들어가게 된다.
- [0039] 본 실시예에 따른 지역 영역 기반의 곱셈 누적 연산 장치는 이전 지역 영역 연산에서 불완전 데이터의 정보를 MAC 모듈 내에 임시 저장하고, 다음 지역 영역이 시작되는 순간 앞서 계산한 결과를 불러오고, 최종 결과물이 온전한 MAC 결과를 낼 수 있도록 한다.
- [0040] 도 6을 참조하면, 지역 영역 2의 회색 부분에 대한 연산 결과를 임시로 저장하고, 다음 지역 영역 3이 연산 될 때 MAC 연산기 내 ACC에 해당 결과를 불러와서 온전한 결과물을 출력한다.
- [0041] 도 7 및 도 8은 본 발명의 일 실시예에 따른 지역 영역 기반의 곱셈 누적 연산 장치를 예시한 도면이다.
- [0042] 지역 영역 기반의 곱셈 누적 연산 장치(10)는 곱셈부(100), 분할부(200), 제1 덧셈부(300), 제2 덧셈부(400), 제어부(500), 후처리부(600)를 포함할 수 있다.
- [0043] 곱셈부(100)는 입력된 제1 데이터와 제2 데이터를 곱셈하여 곱셈 결과를 출력한다.
- [0044] 분할부(200)는 곱셈 결과를 제1 경로와 제2 경로로 분할한다. 분할부(200)는 제1 멀티플렉서와 제2 멀티플렉서를 포함할 수 있다. 교차 연결된 제1 멀티플렉서와 제2 멀티플렉서의 입력으로 곱셈 결과가 입력되거나 초기값 0이 입력될 수 있다.
- [0045] 제1 덧셈부(300)는 제1 경로에 따른 제1 곱셈 결과를 누적하고 저장한다. 제1 덧셈부(300)는 제1 덧셈기와 제1

레지스터를 포함할 수 있다. 제1 레지스터의 값이 제1 덧셈기에 입력되고 제3 멀티플렉서에 입력된다.

- [0046] 제2 덧셈부(400)는 제2 경로에 따른 제2 곱셈 결과를 누적하고 저장한다. 제2 덧셈부(400)는 제2 덧셈기, 제3 멀티플렉서, 제2 레지스터를 포함할 수 있다. 제3 멀티플렉서의 입력으로 제1 레지스터의 값이 입력되고 제2 덧셈기의 값이 입력된다. 제2 레지스터의 값이 제2 덧셈기에 입력된다.
- [0047] 제2 덧셈부(400)는 다음 지역 영역 사이클에서 제1 덧셈부의 누적된 제1 곱셈 결과를 고려할 수 있다.
- [0048] 제어부(500)는 분할부(200), 제1 덧셈부(300), 제2 덧셈부(400)에 제어 신호를 송신하여 동작을 제어하며, 하나의 모듈 또는 복수의 모듈로 구현될 수 있다.
- [0049] 분할부(200), 제1 덧셈부(300), 제2 덧셈부(400)를 통해 입력 데이터가 저장된 메모리에 할당된 수용 영역을 지역 영역으로 분할하여 제1 데이터와 제2 데이터에 대한 연산을 수행할 수 있다.
- [0050] 제어부(500)는 분할부(200)로 제1 선택 신호를 송신하고, 제1 선택 신호에 따라 분할부(200)는 제1 경로 또는 제2 경로를 활성화시킬 수 있다.
- [0051] 제어부(500)는 제1 덧셈부(300) 및 제2 덧셈부(400)로 클록 신호를 송신하고, 클록 신호에 따라 제1 덧셈부(300) 및 제2 덧셈부(400)가 동기화될 수 있다.
- [0052] 제어부(500)는 제1 덧셈부(300)로 리셋 신호를 송신하고, 리셋 신호에 따라 제1 덧셈부(300)는 임시 저장된 데이터를 초기화할 수 있다. 제1 덧셈부(300)는 지역 영역 사이클에 따라 데이터를 임시 저장한다.
- [0053] 제어부(500)는 제2 덧셈부(400)로 제2 선택 신호를 송신하고, 제2 선택 신호에 따라 제2 덧셈부(400)는 제1 덧셈부(300)에 임시 저장된 데이터를 수신할 수 있다.
- [0054] 후처리부(600)는 누적된 제2 곱셈 결과를 가공한다. 후처리부(600)는 네트워크 모델의 연산을 수행하기 위한 활성화 함수 등을 포함할 수 있다.
- [0055] 도 8에서 각 구성요소는 W(Weight data), A(Activation data), X(Multiplier), +(Adder), R(Register), Cnt(Counter), B(Bias Adder), Re(Relu Fx)를 의미한다. 예시된 10, 14, 24는 곱산 누산에 따른 비트수의 변화를 예시한다. R₁은 이전 field의 연산 결과물을 누산 및 임시 저장을 담당하고, R₂는 최종 MAC 연산 결과물을 누산하고, MUX(sel2)를 통해 이전 영역의 데이터를 오프셋으로 사용할 수 있도록 가져오는 역할을 수행한다.
- [0056] 도 9는 본 발명의 일 실시예에 따른 지역 영역 기반의 곱셈 누적 연산 장치가 데이터를 처리하는 타이밍을 예시한 도면이다.
- [0057] 도 9의 타이밍도는 3x3 필터(filter)를 가정한 것으로 이러한 경우 9번의 ACC가 필요하다. 한 번 데이터 처리 단위는 11 클록(clk)이 필요하다. 이러한 주기는 지역 영역 사이클에 대응한다. 지역 영역 사이클에 맞춰 데이터의 연산 결과물을 임시 저장하고 오프셋으로 사용할 수 있다.
- [0058] 도 10 및 도 11은 본 발명의 다른 실시예에 따른 수용 영역의 분할 연산 방법을 예시한 흐름도이다. 수용 영역의 분할 연산 방법은 지역 영역 기반의 곱셈 누적 연산 장치에 의해 수행될 수 있다.
- [0059] 수용 영역의 분할 연산 방법은 입력 데이터가 저장된 메모리에 할당된 수용 영역을 복수의 지역 영역으로 분할하는 단계(S10) 및 복수의 지역 영역에 대해서 연산을 수행하는 단계(S20)를 포함한다.
- [0060] 연산을 수행하는 단계(S20)는, 복수의 지역 영역 중에서 이전 지역 영역에 대해서 연산을 수행한 불완전한 결과 데이터를 임시 저장하는 단계(S21), 복수의 지역 영역 중에서 다음 지역 영역이 시작되는 순간 임시 저장된 불완전한 결과 데이터를 불러오는 단계(S22), 및 임시 저장된 불완전한 결과 데이터를 반영하여 완전한 데이터에 대해서 곱셈 누적 연산을 수행하는 단계(S23)를 포함할 수 있다.
- [0061] 연산을 수행하는 단계(S20)는 이미지의 픽셀 데이터에 대해서 커널 단위로 곱셈 누적 연산을 수행할 수 있다.
- [0062] 연산을 수행하는 단계(S20)는 이미지의 픽셀 데이터에 대해서 초해상도 변환을 수행할 수 있다.
- [0063] 연산을 수행하는 단계(S20)는 학습 네트워크 모델을 통해 입력 처리 단위에 해당하는 수용 영역으로부터 출력 처리 단위에 해당하는 투사 영역으로 연산을 수행하며, 지역 영역 단위를 기준으로 메모리를 접근할 수 있다.
- [0064] 도 12는 본 발명의 실시예들에 따른 시뮬레이션 결과를 예시한 도면이다.
- [0065] 본 실시예들에 따르면 수용 영역을 그대로 적용할 때보다 메모리 बैं드폭을 완화시킬 수 있음을 확인할 수 있다.

이는 수용 영역의 메모리 밴드폭 문제를 완화하여, 시스템에서 더 효율적으로 수용 영역의 크기를 선택할 수 있다. 수용 영역을 크게 가져갈수록 시스템 측면에서 중복되는 연산이 줄게 되어 시스템 성능이 좋아진다. 수용 영역을 한 번에 처리하는 것이 아닌 지역 영역으로 나누어 처리하며, 연산기 내 임시 저장 방식을 통해 구현 가능하다. 제안한 방식을 통해 연산에서 한 번에 접근하는 데이터를 감소시키고, 그에 따라 메모리 밴드폭 문제를 해결할 수 있다. 또한, 유사한 방식으로 CNN 외 수 많은 인공지능 알고리즘에 적용이 가능하다.

[0066] 본 곱셈 누적 연산 장치가 적용된 다양한 전자 장치에 포함된 복수의 구성요소들은 상호 결합되어 적어도 하나의 모듈로 구현될 수 있다. 구성요소들은 장치 내부의 소프트웨어적인 모듈 또는 하드웨어적인 모듈을 연결하는 통신 경로에 연결되어 상호 간에 유기적으로 동작한다. 이러한 구성요소들은 하나 이상의 통신 버스 또는 신호선을 이용하여 통신한다.

[0067] 본 곱셈 누적 연산 장치가 적용된 다양한 전자 장치는 하드웨어, 펌웨어, 소프트웨어 또는 이들의 조합에 의해 로직회로 내에서 구현될 수 있고, 범용 또는 특정 목적 컴퓨터를 이용하여 구현될 수도 있다. 장치는 고정배선형(Hardwired) 기기, 필드 프로그램 가능한 게이트 어레이(Field Programmable Gate Array, FPGA), 주문형 반도체(Application Specific Integrated Circuit, ASIC) 등을 이용하여 구현될 수 있다. 또한, 장치는 하나 이상의 프로세서 및 컨트롤러를 포함한 시스템온칩(System on Chip, SoC)으로 구현될 수 있다.

[0068] 본 곱셈 누적 연산 장치가 적용된 다양한 전자 장치는 하드웨어적 요소가 마련된 컴퓨팅 디바이스에 소프트웨어, 하드웨어, 또는 이들의 조합하는 형태로 탑재될 수 있다. 컴퓨팅 디바이스는 각종 기기 또는 유무선 통신망과 통신을 수행하기 위한 통신 모듈 등의 통신장치, 프로그램을 실행하기 위한 데이터를 저장하는 메모리, 프로그램을 실행하여 연산 및 명령하기 위한 마이크로프로세서 등을 전부 또는 일부 포함한 다양한 장치를 의미할 수 있다.

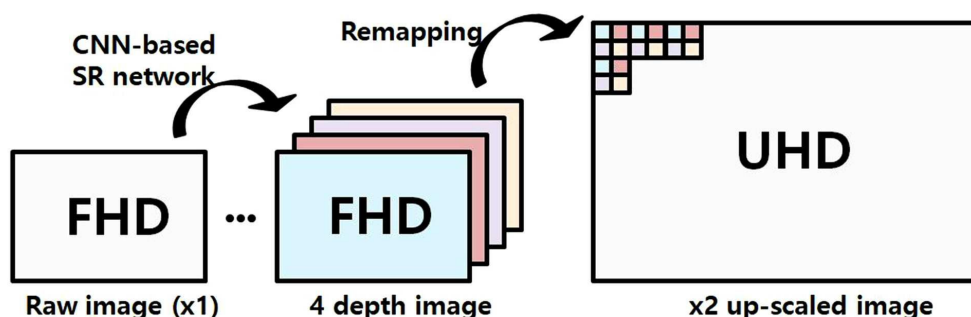
[0069] 도 10 및 도 11에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나 이는 예시적으로 설명한 것에 불과하고, 이 분야의 기술자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 도 10 및 도 11에 기재된 순서를 변경하여 실행하거나 또는 하나 이상의 과정을 병렬적으로 실행하거나 다른 과정을 추가하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이다.

[0070] 본 실시예들에 따른 동작은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능한 매체에 기록될 수 있다. 컴퓨터 판독 가능한 매체는 실행을 위해 프로세서에 명령어를 제공하는 데 참여한 임의의 매체를 나타낸다. 컴퓨터 판독 가능한 매체는 프로그램 명령, 데이터 파일, 데이터 구조 또는 이들의 조합을 포함할 수 있다. 예를 들면, 자기 매체, 광기록 매체, 메모리 등이 있을 수 있다. 컴퓨터 프로그램은 네트워크로 연결된 컴퓨터 시스템 상에 분산되어 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다. 본 실시예를 구현하기 위한 기능적인(Functional) 프로그램, 코드, 및 코드 세그먼트들은 본 실시예가 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있을 것이다.

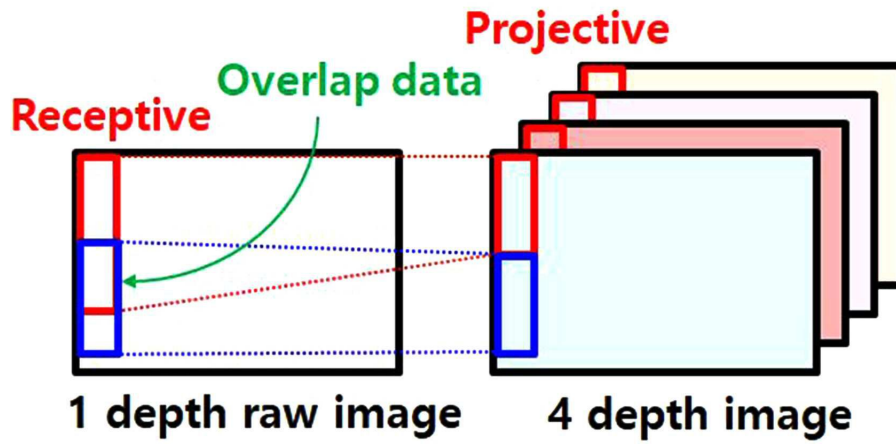
[0071] 본 실시예들은 본 실시예의 기술 사상을 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

도면

도면1



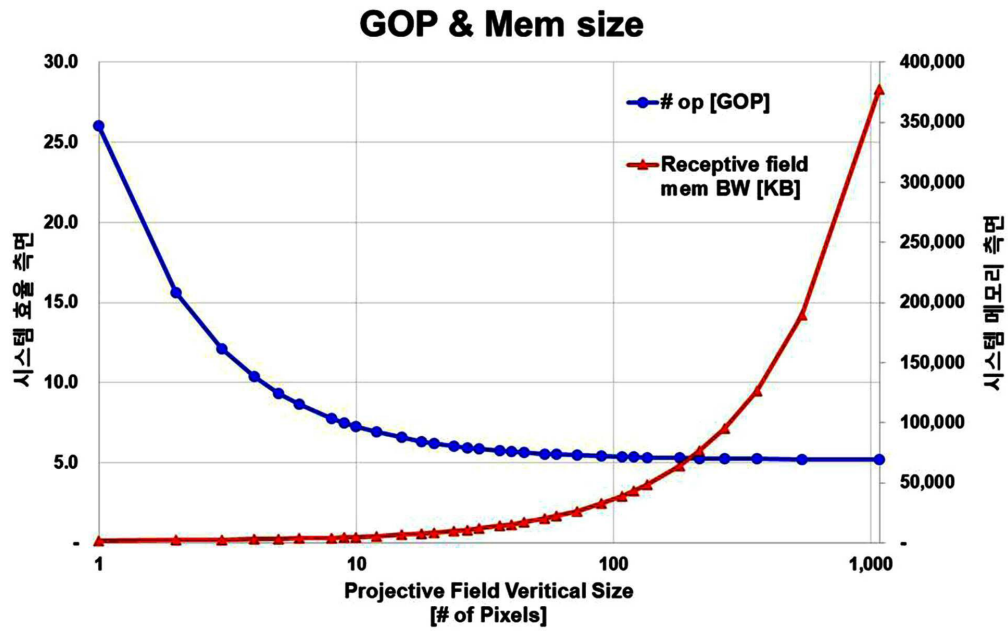
도면2



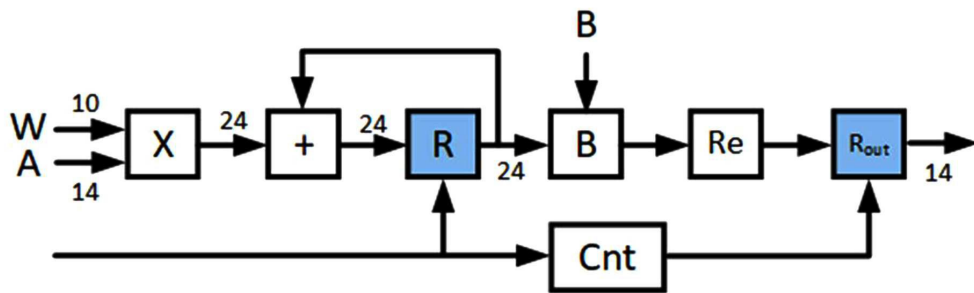
도면3

	전체 이미지	Receptive field	Projective field	반복 횟수
예시 1	1920x 1080	1x9	1x1	1920x 1080
예시 2		1x68	1x60	1920x 18
예시 3		1x80	1x72	1920x 15

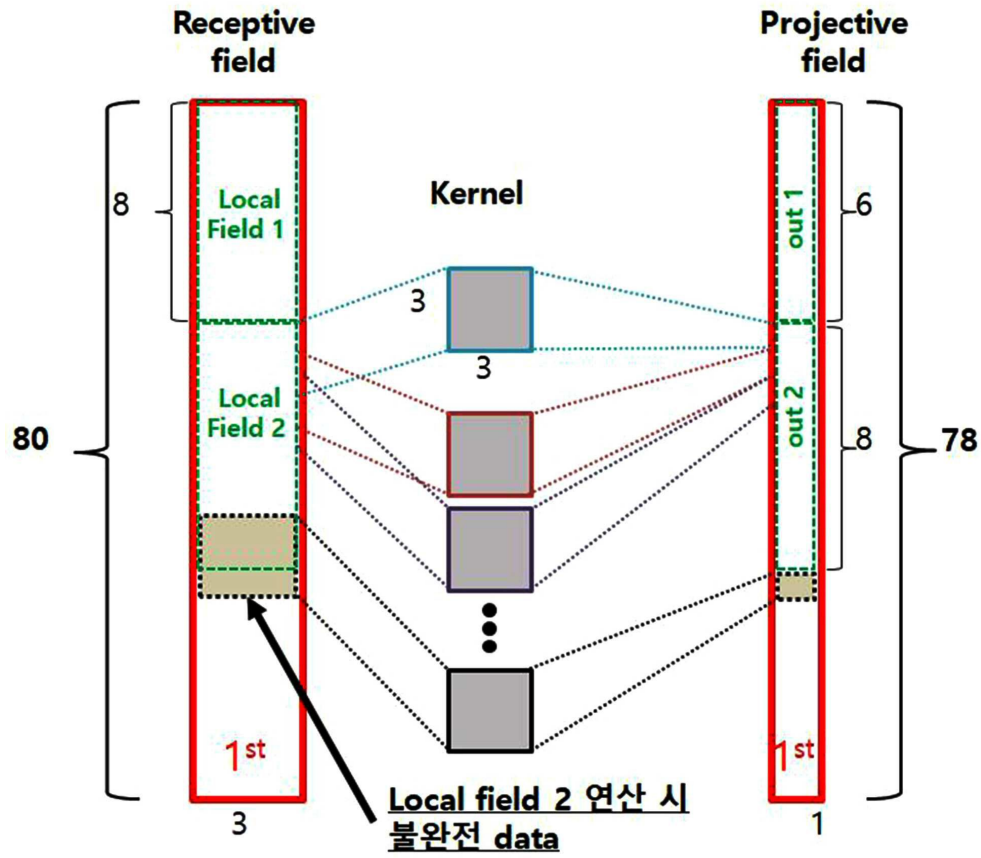
도면4



도면5



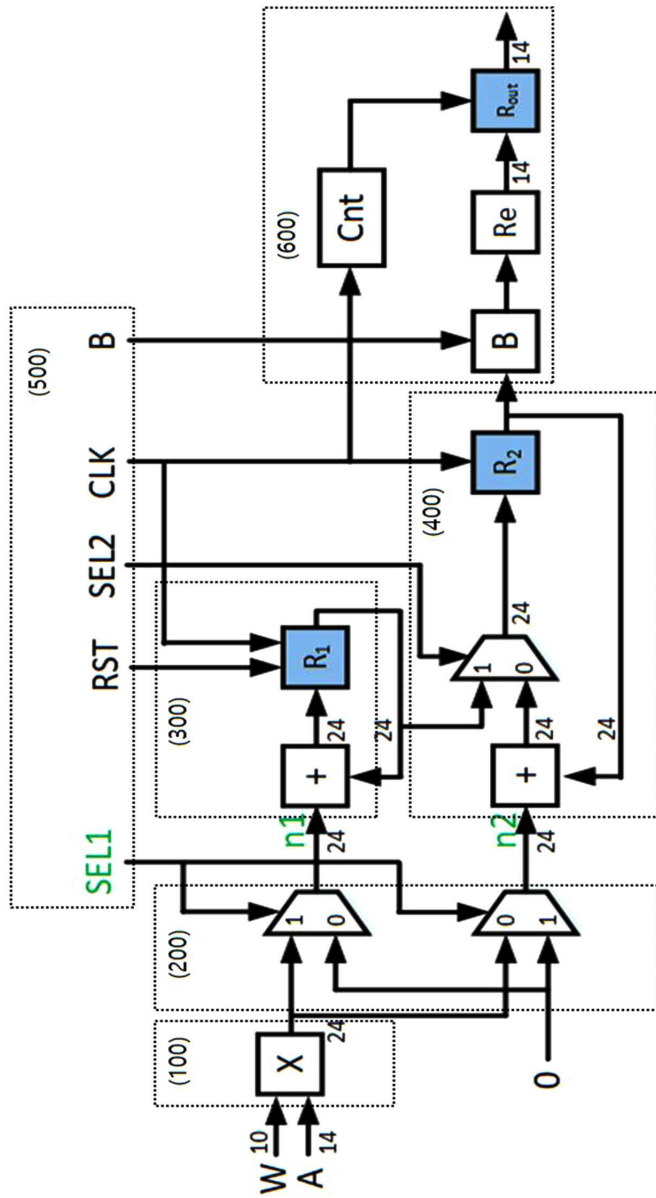
도면6



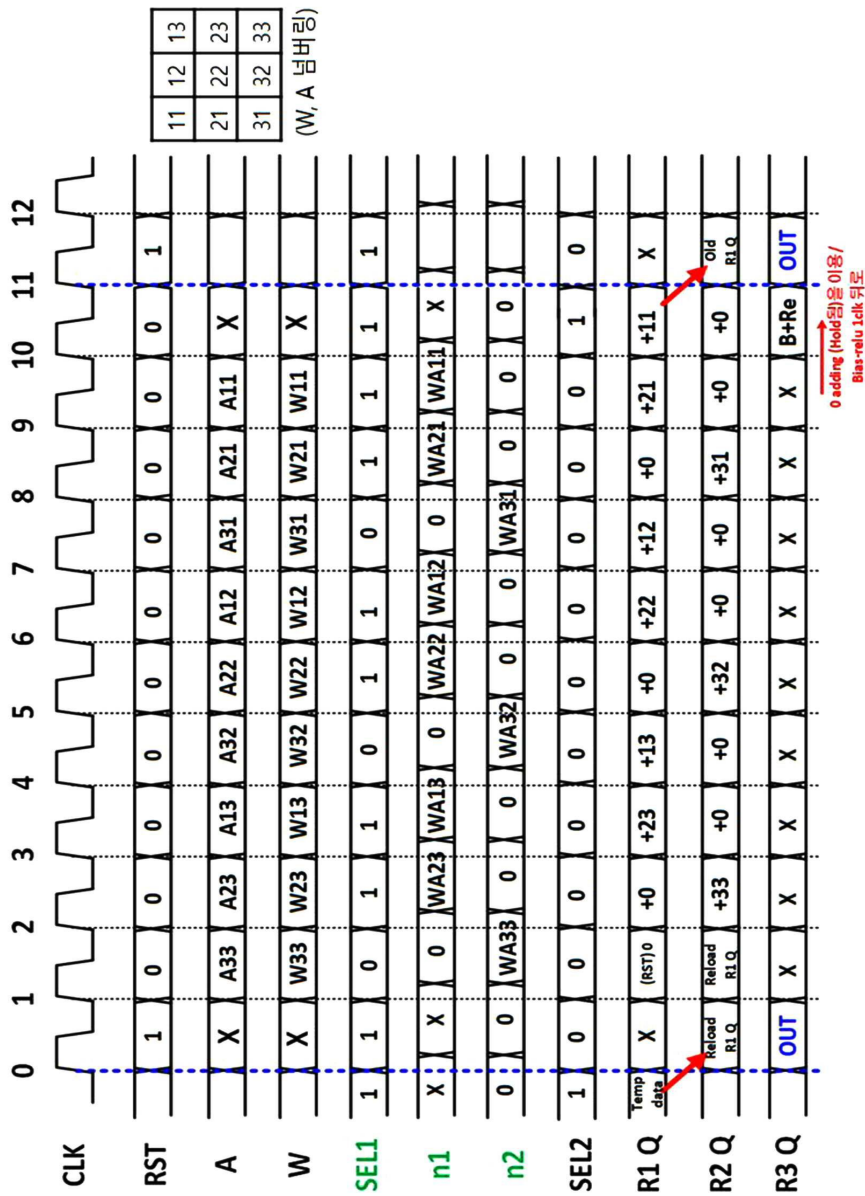
도면7



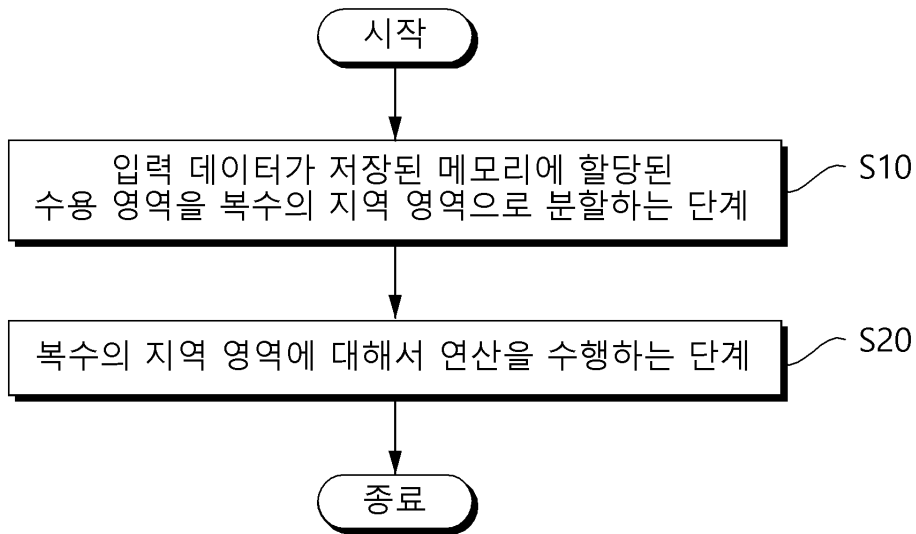
도면8



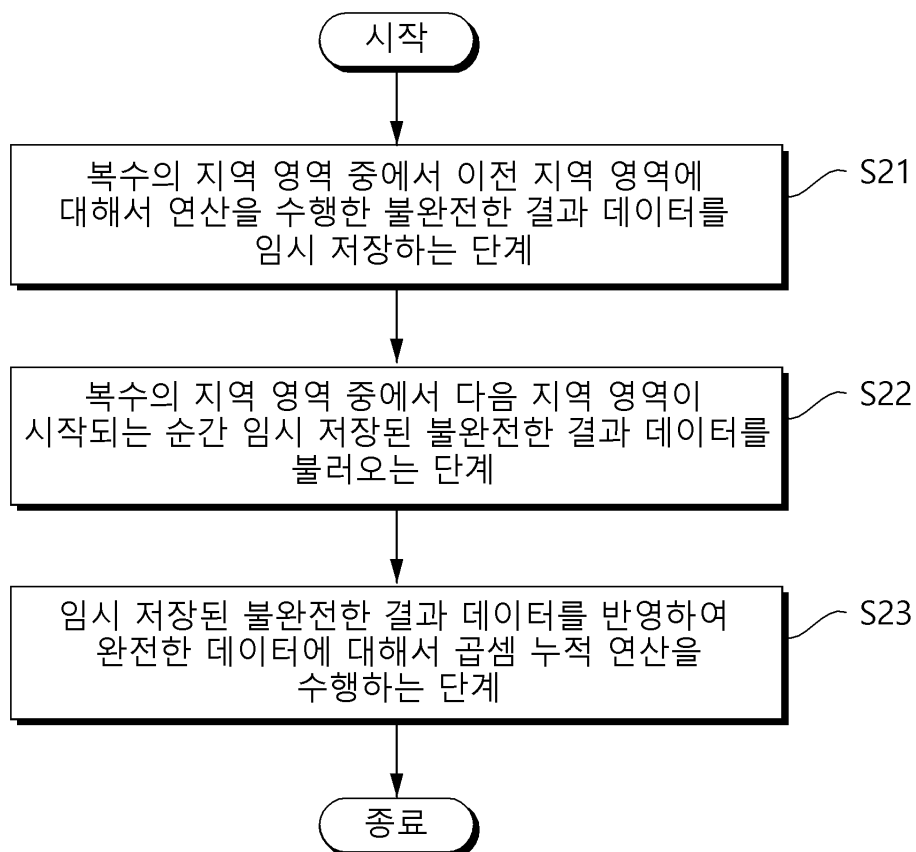
도면9



도면10



도면11



도면12

