



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2023년02월06일

(11) 등록번호 10-2494834

(24) 등록일자 2023년01월30일

(51) 국제특허분류(Int. Cl.)

G16B 45/00 (2019.01) G01N 15/14 (2006.01)

G16B 15/00 (2019.01) G16B 40/20 (2019.01)

(52) CPC특허분류

G16B 45/00 (2019.02)

G01N 15/14 (2013.01)

(21) 출원번호 10-2021-0140815

(22) 출원일자 2021년10월21일

심사청구일자 2021년10월21일

(56) 선행기술조사문헌

W02021075826 A1*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

주식회사 아이티메딕

경기도 용인시 기흥구 흥덕4로30번길 24 ,502호 (영덕동)

(72) 발명자

송재우

서울특별시 서초구 잠원로14길 23, 201동 304호(잠원동, 롯데캐슬갤럭시)

이주범

경기도 수원시 영통구 법조로 38, 101동 203호(하동, 광고 더샵 레이크파크)

(74) 대리인

특허법인우인

전체 청구항 수 : 총 15 항

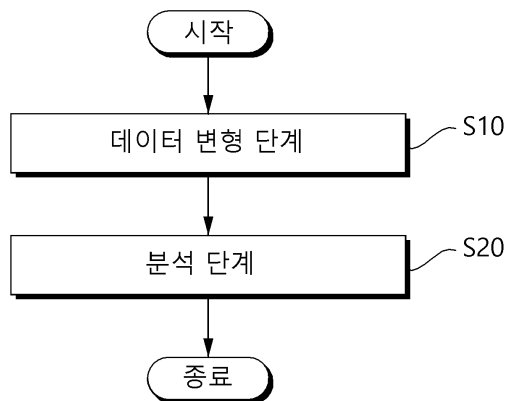
심사관 : 성경아

(54) 발명의 명칭 FCS 데이터 기원 초입방체 형상의 가상 데이터 포인트 추가 및 불완전 데이터 세트 군집 분석을 위한 데이터 처리 장치 및 방법

(57) 요약

본 실시예들은 초기 데이터를 초입방체 공간 내 좌표와 농도 데이터를 갖는 초입방체 형상 데이터로 변환하고, 초입방체 형상 데이터를 테이블 기반의 초입방체 변환 데이터로 표현하고, 초입방체 압축 및 가상 데이터 추가를 통해 그라디언트 군집 분석을 가능하게 하는 데이터 처리 장치 및 방법을 제공한다.

대표도 - 도28



(52) CPC특허분류

G16B 15/00 (2019.02)

G16B 40/20 (2019.02)

G01N 2015/1402 (2013.01)

명세서

청구범위

청구항 1

가상 데이터를 기반으로 초입방체 데이터의 군집을 분석하는 데이터 처리 방법에 있어서,

초기 데이터를 초입방체 공간 내 좌표와 농도 데이터를 갖는 초입방체 형상 데이터로 변환하고, 상기 초입방체 형상 데이터를 테이블 기반의 초입방체 변환 데이터로 표현하는 데이터 변형 단계;

상기 초입방체 형상 데이터 또는 상기 초입방체 변환 데이터를 기반으로 군집 분석을 수행하여 군집 분석 결과를 출력하는 분석 단계를 포함하며,

상기 데이터 변형 단계는 상기 초입방체 형상 데이터 또는 상기 초입방체 변환 데이터에 가상 데이터 포인터를 추가하는 단계를 포함하는 데이터 처리 방법.

청구항 2

제1항에 있어서,

상기 초기 데이터는 혈액의 유세포 분석(Flow Cytometry)에 대한 결과를 나타내는 생체 추출 데이터를 포함하며, 상기 생체 추출 데이터는 미리 설정된 정형화된 형식 또는 FCS(Flow Cytometry Standard) 형식으로 표현되며,

상기 초입방체 변환 데이터는 복수의 축에 해당하는 좌표 정보 및 상기 좌표 정보에 매칭하는 값 정보가 행 방향 또는 열 방향을 기준으로 배치되는 테이블로 표현되는 것을 특징으로 하는 데이터 처리 방법.

청구항 3

제1항에 있어서,

상기 데이터 변형 단계는,

상기 초입방체 형상 데이터 또는 상기 초입방체 변환 데이터를 압축 배수를 기준으로 압축하는 단계를 포함하는 것을 특징으로 하는 데이터 처리 방법.

청구항 4

제3항에 있어서,

상기 압축하는 단계는,

상기 초입방체 변환 데이터에 대해서 압축 배수로 나눈 몫을 기준으로 복수의 라인 병합(merge)을 통해 농도를 증가시키는 것을 특징으로 하는 데이터 처리 방법.

청구항 5

제3항에 있어서,

상기 분석 단계는,

압축된 상기 초입방체 형상 데이터 또는 상기 초입방체 변환 데이터에 대해서 그라디언트 군집 분석을 수행하는 것을 특징으로 하는 데이터 처리 방법.

청구항 6

삭제

청구항 7

제1항에 있어서,

상기 가상 데이터 포인터를 추가하는 단계는,

유효 복셀을 중심으로 하는 초구체를 설정하고,

상기 초구체의 범위 내에 존재하는 유효 복셀과 상기 초구체의 중심에 위치한 유효 복셀의 벡터차를 산출하고,

각 초구체에서 산출한 벡터차를 합하고,

상기 초구체의 중심으로부터 상기 벡터차의 합의 방향과 거리에 위치한 복셀에 포인트 데이터를 추가하는 것을 특징으로 하는 데이터 처리 방법.

청구항 8

제7항에 있어서,

상기 가상 데이터 포인터를 추가하는 단계는,

군집 내부로 향하는 위치 또는 군집 내부에 유효 복셀을 추가하는 과정을 N (상기 N 은 2이상 자연수) 번 반복하여 2^{n-1} 개의 유효 복셀을 추가하는 것을 특징으로 하는 데이터 처리 방법.

청구항 9

제7항에 있어서,

상기 가상 데이터 포인터를 추가하는 단계는,

상기 초구체의 직경을 압축된 초입방체 군집 간 최소 거리 또는 압축 배수로 설정하는 것을 특징으로 하는 데이터 처리 방법.

청구항 10

프로세서 및 상기 프로세서에 의해 실행되는 프로그램을 저장하는 저장 매체를 포함하며, 가상 데이터를 기반으로 초입방체 데이터의 군집을 분석하는 데이터 처리 장치에 있어서,

상기 프로세서는,

초기 데이터를 초입방체 공간 내 좌표와 농도 데이터를 갖는 초입방체 형상 데이터로 변환하고, 상기 초입방체 형상 데이터를 테이블 기반의 초입방체 변환 데이터로 표현하고,

상기 초입방체 형상 데이터 또는 상기 초입방체 변환 데이터를 기반으로 군집 분석을 수행하여 군집 분석 결과를 출력하며,

상기 프로세서는 상기 초입방체 형상 데이터 또는 상기 초입방체 변환 데이터에 가상 데이터 포인터를 추가하는 것을 특징으로 하는 데이터 처리 장치.

청구항 11

제10항에 있어서,

상기 프로세서는,

상기 초입방체 형상 데이터 또는 상기 초입방체 변환 데이터를 압축 배수를 기준으로 압축하는 것을 특징으로 하는 데이터 처리 장치.

청구항 12

제11항에 있어서,

상기 프로세서는,

상기 초입방체 변환 데이터에 대해서 압축 배수로 나눈 몫을 기준으로 복수의 라인 병합(merge)을 통해 농도를 증가시키는 것을 특징으로 하는 데이터 처리 장치.

청구항 13

제11항에 있어서,

상기 프로세서는,

압축된 상기 초입방체 형상 데이터 또는 상기 초입방체 변환 데이터에 대해서 그라디언트 군집 분석을 수행하는 것을 특징으로 하는 데이터 처리 장치.

청구항 14

삭제

청구항 15

제10항에 있어서,

상기 프로세서는,

유효 복셀을 중심으로 하는 초구체를 설정하고,

상기 초구체의 범위 내에 존재하는 유효 복셀과 상기 초구체의 중심에 위치한 유효 복셀의 벡터차를 산출하고,

각 초구체에서 산출한 벡터차를 합하고,

상기 초구체의 중심으로부터 상기 벡터차의 합의 방향과 거리에 위치한 복셀에 포인트 데이터를 추가하는 것을 특징으로 하는 데이터 처리 장치.

청구항 16

제15항에 있어서,

상기 프로세서는,

군집 내부로 향하는 위치 또는 군집 내부에 유효 복셀을 추가하는 과정을 N (상기 N 은 2이상 자연수) 번 반복하여 2^{n-1} 개의 유효 복셀을 추가하는 것을 특징으로 하는 데이터 처리 장치.

청구항 17

제15항에 있어서,

상기 프로세서는,

상기 초구체의 직경을 압축된 초입방체 군집 간 최소 거리 또는 압축 배수로 설정하는 것을 특징으로 하는 데이터 처리 장치.

발명의 설명

기술 분야

[0001] 본 발명이 속하는 기술 분야는 생체 데이터 처리 및 방법에 관한 것이다.

배경 기술

[0002] 도 1은 종래의 FCS 데이터 분석 방식의 일례를 설명하기 위한 도면이다.

[0003] 도 1을 참조하면, 종래의 FCS 데이터 분석 방식은 FCS(flow cytometry standard) 데이터를 히스토그램 또는 좌표 평면의 점도표로 표현하고, 육안으로 점들의 군집을 인식하여 수기 또는 반자동으로 군집을 구분하는 경계선을 설정하는 방법을 취하고 있다. 이는 FCS 데이터 분석의 가장 기본적인 분석 방식이며 또한 널리 사용되는 기법이다.

[0004] 도 2는 종래의 FCS 데이터 분석 방식의 다른 예를 설명하기 위한 도면이다.

[0005] 도 2를 참조하면, SOM(self organizing map)은 저차원(예컨대, 2차원 등) 공간 내 격자를 따라 배치된 점들에

입력 데이터와 같은 차원의 벡터를 적용하고, 각 점들의 위치를 데이터 점들의 분포 위치로 변경하면서 유클리드 거리(euclidean distance)의 합이 최소가 되도록 데이터 점들을 각각의 격자점들에 배치하는 기계학습 방법이다. 분석 결과로 고차원 데이터 각 개체들의 군집을 구분하여 저차원 형식으로 표현한다. 즉, 고차원의 데이터 원 공간에 분포하는 유사한 개체들은 저차원에 인접한 격자점들과 연결되는 것이다. 각 격자 점은 하나의 군집 역할을 한다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) KR 10-1857624 (2018.05.08)

발명의 내용

해결하려는 과제

[0007] 본 발명의 실시예들은 초기 데이터를 초입방체 공간 내 좌표와 농도 데이터를 갖는 초입방체 형상 데이터로 변환하고, 초입방체 형상 데이터를 테이블 기반의 초입방체 변환 데이터로 표현하고, 초입방체 압축 및 가상 데이터 추가를 통해 그라디언트 군집 분석이 가능하도록 데이터를 처리하는데 주된 목적이 있다.

[0008] 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

[0009] 본 실시예의 일 측면에 의하면 데이터 처리 방법에 있어서, 초기 데이터를 초입방체 공간 내 좌표와 농도 데이터를 갖는 초입방체 형상 데이터로 변환하고, 상기 초입방체 형상 데이터를 테이블 기반의 초입방체 변환 데이터로 표현하는 데이터 변형 단계; 상기 초입방체 형상 데이터 또는 상기 초입방체 변환 데이터를 기반으로 분석을 수행하여 분석 결과를 출력하는 분석 단계를 포함하는 데이터 처리 방법을 제공한다.

[0010] 본 실시예의 다른 측면에 의하면 프로세서 및 상기 프로세서에 의해 실행되는 프로그램을 저장하는 저장 매체를 포함하는 데이터 처리 장치에 있어서, 상기 프로세서는, 초기 데이터를 초입방체 공간 내 좌표와 농도 데이터를 갖는 초입방체 형상 데이터로 변환하고, 상기 초입방체 형상 데이터를 테이블 기반의 초입방체 변환 데이터로 표현하고, 상기 초입방체 형상 데이터 또는 상기 초입방체 변환 데이터를 기반으로 분석을 수행하여 분석 결과를 출력하는 것을 특징으로 하는 데이터 처리 장치를 제공한다.

발명의 효과

[0011] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 초기 데이터를 초입방체 공간 내 좌표와 농도 데이터를 갖는 초입방체 형상 데이터로 변환하고, 초입방체 형상 데이터를 테이블 기반의 초입방체 변환 데이터로 표현하고, 초입방체 압축 및 가상 데이터 추가를 통해 그라디언트 군집 분석을 가능하게 하는 효과가 있다.

[0012] 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

[0013] 도 1은 종래의 FCS 데이터 분석 방식의 일례를 설명하기 위한 도면이다.

도 2는 종래의 FCS 데이터 분석 방식의 다른 예를 설명하기 위한 도면이다.

도 3은 본 발명의 일 실시예에 따른 데이터 처리 장치를 예시한 도면이다.

도 4는 본 발명의 일 실시예에 따른 FCS 데이터의 초입방체 형상 테이블을 기초로 초입방체 픽셀의 그라디언트 분석을 통해 자동으로 군집을 분석하는 장치를 설명하기 위한 블록도이다.

도 5는 일반 혈액 검사(complete blood cell count, CBC)의 일례를 설명하기 위한 도면이다.

- 도 6은 본 발명의 일 실시예에 따른 FCS 데이터의 일례를 설명하기 위한 도면이다.
- 도 7은 본 발명의 일 실시예에 따른 초입방체 형상 전환 과정을 설명하기 위한 도면이다.
- 도 8은 본 발명의 일 실시예에 따른 자동 혈구 분석기를 통한 백혈구 분석 채널의 형상 전환 일례를 설명하기 위한 도면이다.
- 도 9는 본 발명의 일 실시예에 따른 초입방체 형상 테이블의 일례를 설명하기 위한 도면이다.
- 도 10은 본 발명의 일 실시예에 따른 초입방체 변환 데이터의 일례를 설명하기 위한 도면이다.
- 도 11은 본 발명의 일 실시예에 따른 초입방체 그레이 스케일 테이블의 일례를 설명하기 위한 도면이다.
- 도 12는 본 발명의 일 실시예에 따른 초입방체 그라디언트 스케일 테이블의 일례를 설명하기 위한 도면이다.
- 도 13은 FCS 데이터를 히스토그램으로 나타낸 예시적인 도면이다.
- 도 14는 FCS 데이터를 입방체로 나타낸 예시적인 도면이다.
- 도 15는 본 발명의 일 실시예에 따른 자동 군집 분석 과정을 설명하기 위한 도면이다.
- 도 16은 본 발명의 일 실시예에 따른 자동 군집 분석 과정에 이용되는 거리 계산 방법의 일례를 설명하기 위한 도면이다.
- 도 17은 본 발명의 일 실시예에 따른 복셀 상태의 일례를 설명하기 위한 도면이다.
- 도 18 및 도 19는 본 발명의 일 실시예에 따른 압축된 초입방체 변환 데이터의 일례를 설명하기 위한 도면이다.
- 도 20은 본 발명의 일 실시예에 따른 압축된 초입방체 형상 데이터의 일례를 설명하기 위한 도면이다.
- 도 21은 본 발명의 일 실시예에 따른 초구체 설정 동작을 설명하기 위한 도면이다.
- 도 22는 본 발명의 일 실시예에 따른 벡터차의 방향을 설명하기 위한 도면이다.
- 도 23은 본 발명의 일 실시예에 따른 포인트 데이터 추가 동작을 설명하기 위한 도면이다.
- 도 24는 본 발명의 일 실시예에 따른 포인트 데이터 추가로 인한 그라디언트 기법 적용시 군집 상태 변화를 설명하기 위한 도면이다.
- 도 25는 본 발명의 일 실시예에 따른 초구체 직경과 군집 상태의 관계를 설명하기 위한 도면이다.
- 도 26은 본 발명의 일 실시예에 따른 최소 군집 간 거리를 초구체 직경으로 사용한 결과를 설명하기 위한 도면이다.
- 도 27은 본 발명의 일 실시예에 따른 질병 판단 모델의 기계학습에 이용되는 학습 데이터의 일례를 설명하기 위한 도면이다.
- 도 28 내지 도 31은 본 발명의 다른 실시예에 따른 데이터 처리 방법을 예시한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0014] 이하, 첨부된 도면을 참조하여 본 발명의 실시 예를 상세히 설명한다. 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시 예를 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 게시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 게시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.
- [0015] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어(기술 및 과학적 용어를 포함)는 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있을 것이다. 또 일반적으로 사용되는 사전에 정의되어 있는 용어들은 명백하게 특별히 정의되어 있지 않는 한 이상적으로 또는 과도하게 해석되지 않는다.
- [0016] 본 명세서에서 "제1", "제2" 등의 용어는 하나의 구성요소를 다른 구성요소로부터 구별하기 위한 것으로, 이들 용어들에 의해 권리범위가 한정되어서는 아니 된다. 예를 들어, 제1 구성요소는 제2 구성요소로 명명될 수

있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다.

- [0017] 본 명세서에서 각 단계들에 있어 식별부호(예를 들어, a, b, c 등)는 설명의 편의를 위하여 사용되는 것으로 식별부호는 각 단계들의 순서를 설명하는 것이 아니며, 각 단계들은 문맥상 명백하게 특정 순서를 기재하지 않는 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 단계들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.
- [0018] 본 명세서에서, "가진다", "가질 수 있다", "포함한다" 또는 "포함할 수 있다"등의 표현은 해당 특징(예: 수치, 기능, 동작, 또는 부품 등의 구성요소)의 존재를 가리키며, 추가적인 특징의 존재를 배제하지 않는다.
- [0019] 또한, 본 명세서에 기재된 '~부'라는 용어는 소프트웨어 또는 FPGA(field-programmable gate array) 또는 ASIC과 같은 하드웨어 구성요소를 의미하며, '~부'는 어떤 역할들을 수행한다. 그렇지만 '~부'는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. '~부'는 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 '~부'는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로코드, 회로, 데이터 구조들 및 변수들을 포함한다. 구성요소들과 '~부'들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 '~부'들로 결합되거나 추가적인 구성요소들과 '~부'들로 더 분리될 수 있다.
- [0020] 이하에서 첨부한 도면을 참조하여 본 발명에 따른 FCS 데이터의 초입방체 데이터를 기초로 초입방체 픽셀의 그라디언트 분석을 통해 군집을 분석하는 장치 및 방법의 일 실시예에 대해 상세하게 설명한다.
- [0021] 도 3은 본 발명의 일 실시예에 따른 데이터 처리 장치를 예시한 도면이다.
- [0022] 장치(11)는 적어도 하나의 프로세서(120), 컴퓨터 판독 가능한 저장매체(13) 및 통신 버스(17)를 포함한다.
- [0023] 프로세서(120)는 장치(11)로 동작하도록 제어할 수 있다. 예컨대, 프로세서(12)는 컴퓨터 판독 가능한 저장 매체(130)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 컴퓨터 실행 가능 명령어는 프로세서(12)에 의해 실행되는 경우 장치(11)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.
- [0024] 컴퓨터 판독 가능한 저장 매체(13)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보는 입출력 인터페이스(15)나 통신 인터페이스(16)를 통해서도 주어질 수 있다. 컴퓨터 판독 가능한 저장 매체(13)에 저장된 프로그램(14)은 프로세서(12)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독 가능한 저장 매체(13)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 데이터 처리 장치(11)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.
- [0025] 통신 버스(17)는 프로세서(12), 컴퓨터 판독 가능한 저장 매체(13)를 포함하여 데이터 처리 장치(11)의 다른 다양한 컴포넌트들을 상호 연결한다.
- [0026] 장치(11)는 또한 하나 이상의 입출력 장치를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(15) 및 하나 이상의 통신 인터페이스(16)를 포함할 수 있다. 입출력 인터페이스(15) 및 통신 인터페이스(16)는 통신 버스(17)에 연결된다. 입출력 장치(미도시)는 입출력 인터페이스(15)를 통해 장치(11)의 다른 컴포넌트들에 연결될 수 있다.
- [0027] 프로세서는, 초기 데이터를 초입방체 공간 내 좌표와 농도 데이터를 갖는 초입방체 형상 데이터로 변환하고, 초입방체 형상 데이터를 테이블 기반의 초입방체 변환 데이터로 표현하고, 초입방체 형상 데이터 또는 초입방체 변환 데이터를 기반으로 분석을 수행하여 분석 결과를 출력한다.
- [0028] 초기 데이터는 혈액의 유세포 분석(Flow Cytometry)에 대한 결과를 나타내는 생체 추출 데이터를 포함하며, 생체 추출 데이터는 미리 설정된 정형화된 형식 또는 FCS(Flow Cytometry Standard) 형식으로 표현될 수 있다.
- [0029] 초입방체 변환 데이터는 복수의 축에 해당하는 좌표 정보 및 좌표 정보에 매칭하는 값 정보가 행 방향 또는 열 방향을 기준으로 배치되는 테이블로 표현될 수 있다.
- [0030] 프로세서는 초입방체 형상 데이터 또는 초입방체 변환 데이터를 압축 배수를 기준으로 압축할 수 있다. 프로세

서는 초입방체 변환 데이터에 대해서 압축 배수로 나눈 몫을 기준으로 복수의 라인 병합(merge)을 통해 농도를 증가시킬 수 있다.

- [0031] 프로세서는 압축된 초입방체 형상 데이터 또는 압축된 초입방체 변환 데이터에 대해서 그라디언트 군집 분석을 수행할 수 있다.
- [0032] 프로세서는 초입방체 형상 데이터 또는 초입방체 변환 데이터에 가상 데이터 포인터를 추가할 수 있다. 프로세서는 압축된 초입방체 형상 데이터 또는 압축된 초입방체 변환 데이터에 가상 데이터 포인터를 추가할 수 있다.
- [0033] 프로세서는 유효 복셀을 중심으로 하는 초구체를 설정하고, 초구체의 범위 내에 존재하는 유효 복셀과 상기 초구체의 중심에 위치한 유효 복셀의 벡터차를 산출하고, 각 초구체에서 산출한 벡터차를 합하고, 초구체의 중심으로부터 벡터차의 합의 방향과 거리에 위치한 복셀에 포인트 데이터를 추가할 수 있다.
- [0034] 프로세서는 군집 내부로 향하는 위치 또는 군집 내부에 유효 복셀을 추가하는 과정을 $N(N \geq 2 \text{ 이상 자연수})$ 번 반복하여 2^{n-1} 개의 유효 복셀을 추가할 수 있다.
- [0035] 프로세서는 초구체의 직경을 압축된 초입방체 군집 간 최소 거리 또는 압축 배수로 설정할 수 있다.
- [0036] 도 4는 본 발명의 일 실시예에 따른 FCS 데이터의 초입방체 형상 테이블을 기초로 초입방체 픽셀의 그라디언트 분석을 통해 자동으로 군집을 분석하는 장치를 설명하기 위한 블록도이다.
- [0037] 도 4을 참조하면, 본 발명의 일 실시예에 따른 FCS 데이터의 초입방체 형상 테이블을 기초로 초입방체 픽셀의 그라디언트 분석을 통해 자동으로 군집을 분석하는 장치(이하 '자동 군집 분석 장치'라 한다)(100)는 FCS(flow cytometry standard) 데이터를 기반으로 초입방체(hypercube) 형상 테이블을 획득하고, 초입방체 형상 테이블의 성분별 그라디언트(gradient)를 분석을 통해 자동으로 군집을 분석하여 FCS 데이터에 대응되는 군집을 획득한다.
- [0038] 보다 자세히 설명하면, FCS 데이터는 수집 시간 순으로 배열한 n 차원(n 개의 파라미터) 입력 벡터들의 집합으로 구성된다. 데이터 개체들은 각 벡터들 간의 거리에 따라 유사한 벡터끼리 군집을 형성한다. 이를 초입방체 내 형상으로 전환할 경우 무수히 많은 점들이 통계적 분포를 보이며 군집을 형성하게 된다. 각 군집 내 점들의 통계적 분포에 따라 특정 좌표에서 최대 점 밀도를 보이며 이를 그 군집의 대표점으로 설정한다.
- [0039] 초입방체 내 점 분포밀도를 초입방체를 구성하는 각 단위 공간(픽셀) 별 그레이 스케일 강도(gray scale intensity)로 표현할 경우 군집의 대표점은 최대 강도(intensity)를 보이므로 강도(intensity) 변화율 또는 그라디언트(gradient)가 0이 된다는 점에 착안하여 본 발명에 따른 군집 분석 방법을 고안한다.
- [0040] 본 발명은 초입방체를 구성하는 각 픽셀의 그라디언트(즉, 파라미터 성분별 기울기)를 계산 수집하여 자동으로 군집의 대표점(위치) 및 범위 설정을 수행한다. 본 발명에 따른 군집 분석은 n 차원 벡터들의 유클리드 거리로 표현되는 유사성에 따라 군집을 배정하는 SOM(self organizing map) 방식이 아닌, 고전적인 군집 인식 방식과 유사성을 띠는 직관적 성격의 자동 분석 방법이다. 군집 분석 결과 및 각 군집의 특성들로부터 다양한 질병들의 특징을 찾아내고, 그러한 특징들과 질병과의 연관성을 파악하여 임상적 예측 모델을 고안하는데 활용할 수 있다.
- [0041] 이를 위해, 자동 군집 분석 장치(100)는 데이터 변형부(110), 그라디언트 분석부(130), 군집 분석부(150) 및 질병 판단부(170)를 포함할 수 있다.
- [0042] 데이터 변형부(110)는 FCS 데이터를 기반으로 각 입자의 파라미터 측정값들을 이용하여 각 입자를 각 파라미터를 축으로 하고 파라미터의 측정값들을 좌표 위치로 하는 초입방체의 점(dot)으로 표현하는 초입방체 형상 테이블을 획득한다.
- [0043] 데이터 변형부(110)는 FCS 데이터의 파라미터 측정값들을 각 입자별로 병합하여 각 입자의 좌표 위치를 획득할 수 있다. 이때, 데이터 변형부(110)는 FCS 데이터의 전체 파라미터 중에서 적어도 2개 이상의 파라미터에 대한 측정값들을 각 입자별로 병합하여 각 입자의 좌표 위치를 획득할 수 있다. 또한, 데이터 변형부(110)는 FCS 데이터에 포함된 전체 또는 일부 파라미터의 측정값들을 미리 선정된 일정한 값(예컨대, 4, 8, 32 등의 특정 값)으로 나눈 몫으로 치환하는 방법으로 데이터를 변형(이미지 깊이 변환)하고, 이때 발생하는 데이터의 유실을 막기 위해 각 몫에 소정의 값(예컨대, 10 등)을 더하는 방법을 취할 수 있다.
- [0044] 데이터 변형부(110)는 획득한 각 입자의 좌표 위치를 크기순으로 재배열할 수 있다.

- [0045] 데이터 변형부(110)는 좌표 위치에 대응되는 그레이 스케일 값을 해당 좌표 위치를 가지는 입자의 개수를 기초로 획득하여, 좌표 위치 항목과 그레이 스케일 항목으로 이루어지는 초입방체 형상 테이블을 획득할 수 있다. 예컨대, 변형된 데이터에 포함된 좌표값과 동일한 좌표값이 존재하는 경우, 동일한 좌표값을 삭제하고 해당 좌표값에 대한 카운트값을 기 설정된 단위(예컨대, 1 등)로 증가하여 카운트값을 갱신하고, 변형된 데이터와 갱신된 카운트값을 포함하는 데이터 테이블을 생성한다. 예를 들어, 변형된 데이터의 좌표값이 1 개일 경우 카운트값을 1로 부여하고, 동일한 좌표값이 존재하는 경우 해당 좌표값의 카운트값을 2로 부여하는 방식으로 새로운 데이터 테이블을 생성할 수 있다.
- [0046] 도 5는 일반 혈액 검사(complete blood cell count, CBC)의 일례를 설명하기 위한 도면이고, 도 6은 본 발명의 일 실시예에 따른 FCS 데이터의 일례를 설명하기 위한 도면이고, 도 7은 본 발명의 일 실시예에 따른 초입방체 형상 전환 과정을 설명하기 위한 도면이고, 도 8은 본 발명의 일 실시예에 따른 자동 혈구 분석기를 통한 백혈구 분석 채널의 형상 전환 일례를 설명하기 위한 도면이고, 도 9는 본 발명의 일 실시예에 따른 초입방체 변환 데이터의 일례를 설명하기 위한 도면이고, 도 10은 본 발명의 일 실시예에 따른 초입방체 형상 테이블의 일례를 설명하기 위한 도면이다.
- [0047] 보다 자세히 설명하면, 본 발명의 입력 데이터는 FCS 형식 표준을 따르며 군집 분석의 대상이 되는 모든 데이터를 포함한다. FCS는 유세포 분석기 또는 이와 유사한 분석 원리를 사용하는 검사 장비에서 유래하는 원 데이터의 표준 표기법이며, 분석 대상인 세포 또는 이에 준하는 입자 각각의 파라미터 정보(파라미터별 신호 강도)로 구성된 집합을 데이터의 내용(content)으로 한다.
- [0048] FCS 파일의 헤더(header)로부터 파라미터의 수와 각 파라미터별 채널(channel)에 대한 정보를 취하고 FCS 파일의 내용(content)은 전환 알고리즘을 통해 초입방체 내 점 분포 형상 또는 이를 나타내는 데이터 테이블로 전환된다.
- [0049] FCS 데이터의 군집 분석은 유사한 파라미터 신호 프로파일(profile)을 보이는 세포 및 입자들을 동일한 아집합(subset) 또는 아군(subgroup)으로 분류하고 그 아집합(아군)의 크기를 분석함으로써, 도 5에 도시된 바와 같이, 시료 내 세포/입자의 성분 별 구성 및 분포를 분석하는 것을 목적으로 한다.
- [0050] 초기 입력 데이터는 검체/환자/분석 별로 하나의 파일에 포함되어 있는 FCS 데이터 값들로 구성되어 있다. 데이터의 구성은 각 분석 항목(파라미터)을 열에 나열하고 각 분석 세포/입자들의 파라미터별 측정값을 행별로 나열하는 형식으로 되어 있다.
- [0051] 도 6은 각 세포/입자별로 4개의 파라미터(AAA, BBB, CCC, DDD)의 측정값들을 시간순으로 나열해 놓은 FCS 데이터의 예시이다. 이로부터 각 파라미터를 축으로 하는 4차원 입방체를 구상하고 그 측정값들을 좌표 위치로 생각하면, 각 세포/입자를 하나의 점으로 표시할 수 있으며, 데이터에 포함된 전체 세포/입자를 표시하면 이는 4차원 입방체 내 4차원 그레이 스케일 형상으로 표시할 수 있다. 실제로 4차원 이상의 형상은 그림으로 나타낼 수 없기 때문에, 4개의 파라미터 중 3개를 골라 3차원적 공간의 좌표점들로 표현하면 도 7과 같다. 도 7은 3차원 입방체 내 그레이 스케일 형상이며, 이는 2차원 이미지가 기계 학습(CNN 등)에 사용되는 것과 동일한 방식으로 기계 학습의 입력 데이터로 사용될 수 있다.
- [0052] 이와 같은 n-차원(n-파라미터) FCS 데이터가 형상화되었을 때 환자/검체 간에 보이는 유사성과 미세한 차이를 분석하여 임상적으로 의미 있는 형태적 특성을 추출할 수 있다. 예컨대, 도 8에 도시된 바와 같이, 자동 혈구 분석 검사 유래 FCS 데이터의 형상 전환 데이터로부터 환자/검체 간 유사성과 차이점을 확인할 수 있다.
- [0053] 그리고, 시간순으로 나열된 FCS 데이터의 n-파라미터 정보를 자동으로 n-차원 초입방체 내 그레이 스케일 형상으로 전환한다. 예컨대, 각 파라미터가 k-채널의 측정 범위를 갖는 조건 하에, 각 열별 4개의 파라미터에 대한 FCS 측정값들을 단순히 병기(concatanation)하면, n열의 테이블은 n자리수를 가지는 k진수 수들의 어레이로 단 순화된다. 이 수들을 크기순으로 재배열하면, n차원 초입방체 내 각 위치(픽셀)에 몇개의 데이터 개체들이 배정되는지 쉽게 파악하고, 이를 그레이 스케일로 전환하여 형상화할 수 있다. 도 9에 도시된 초입방체 형상 테이블은 도 7에 도시한 FCS 데이터를 4차원 초입방체 내 픽셀 정보(그레이 스케일)를 나타내는 2열의 테이블로 전환한 예시이다. 좌표값이 1개 일때는 카운트(count)값이 1이 되고, 동일한 좌표값이 2개일때는 카운트(count)값이 2가 된다.
- [0054] 예컨대, 이차원 이미지 파일은 각 픽셀의 위치를 나타내는 좌표를 두 열에 표시하고 나머지 한 열에 각 픽셀의 그레이 스케일 또는 RGB 강도(intensity)를 표시한 테이블 형식을 띤다. 이와 같이 초입방체 형상을 테이블로 표시하자면 각 파라미터 별로 열을 만들고 각 파라미터의 스케일된 신호 강도(scaled signal intensity)가 픽셀

의 좌표를 나타내도록 하면 된다. 두 입자가 동일한 파라미터 값을 보이는 경우 이들은 초입방체 내 같은 픽셀에 로컬리제이션(localization)되며, 이로 인해 이 픽셀은 2 그레이 스케일로 표시된다. 마찬가지로 네 입자가 동일한 파라미터를 보일 경우 이 파라미터 프로파일에 해당하는 위치의 픽셀은 4 그레이 스케일로 표시된다. 유세포 분석 후 얻은 FCS 데이터의 많은 입자(또는 세포)들을 이러한 방식으로 각각 (n차) 초입방 픽셀에 로컬리제이션(localization)시키면 (n차) 초입방체 내 음영 형상이 형성된다. 이 형상 파일은 앞서 설명한 형상데이터블의 형식을 갖는다. 이로써 시간 순으로 배열된 입자들의 n-파라미터 데이터가 초입방체를 형성하는 데이터로 전환되게 된다. 이러한 방식의 데이터 전환은 현재 멀티-파라미터 데이터의 군집 분석에 사용되는 SOM 방식에 비해, 기존의 육안 확인을 통해 수기로 군집을 인식하는 직관적 방식의 군집 분석을 가능하게 한다.

[0055] 도 10을 참조하면, 파라미터 수가 4 개 이상의 멀티파라미터 FCS 데이터를 이와 같이 직접 시각적으로 표현할 수는 없지만 개념적 초입방체 공간 내 복셀(voxel)의 좌표 및 그레이스케일 밀도(grayscale intensity) 테이블로 표현할 수 있다.

[0056] 그라디언트 분석부(130)는 데이터 변형부(110)에 의해 획득된 초입방체 형상 테이블을 기반으로 각 점의 좌표 위치를 파라미터별로 분리하여 표현하는 초입방체 그레이 스케일(gray scale) 테이블을 획득하고, 획득된 초입방체 그레이 스케일 테이블을 기반으로 각 점별로 그라디언트를 분석하여 각 점별 그라디언트 값을 포함하는 초입방체 그라디언트 스케일 테이블을 획득한다.

[0057] 그라디언트 분석부(130)는 초입방체 형상 테이블의 좌표 위치 항목을 파라미터별로 분리하여 초입방체 그레이 스케일 테이블을 획득할 수 있다.

[0058] 그라디언트 분석부(130)는 획득된 초입방체 그레이 스케일 테이블의 각 좌표 위치 항목에 대해 파라미터별로 그라디언트를 분석하여 파라미터별 그라디언트 값을 획득하여, 복수개의 파라미터별로 분리되어 있는 좌표 위치 항목, 그레이 스케일 항목 및 그라디언트 항목으로 이루어지는 초입방체 그라디언트 스케일 테이블을 획득할 수 있다.

[0059] 그라디언트 분석부(130)는 초입방체 그레이 스케일 테이블의 각 좌표 위치 항목에 대해, 해당 좌표 위치 항목에 대응되는 그레이 스케일 값과 파라미터별로 인접한 복수개의 좌표 위치 항목에 대응되는 복수개의 그레이 스케일 값을 이용하여 해당 좌표 위치 항목에 대한 파라미터별 그라디언트 값을 획득할 수 있다. 예컨대, 좌표 위치 항목이 "제1 파라미터", "제2 파라미터" 및 "제3 파라미터"의 3개의 파라미터별로 분리된 경우, 특정 좌표 위치에 대응되는 그라디언트 값은 파라미터의 개수인 "3개"로 이루어질 수 있다. 즉, 특정 좌표 위치에서 "제1 파라미터" 방향으로 인접한 좌표 위치의 그레이 스케일 값, 특정 좌표 위치에서 "제2 파라미터" 방향으로 인접한 좌표 위치의 그레이 스케일 값 및 특정 좌표 위치에서 "제3 파라미터" 방향으로 인접한 좌표 위치의 그레이 스케일 값 각각에서 특정 좌표 위치의 그레이 스케일 값을 빼서, 파라미터별 그라디언트 값을 획득할 수 있다.

[0060] 도 11은 본 발명의 일 실시예에 따른 초입방체 그레이 스케일 테이블의 일례를 설명하기 위한 도면이고, 도 12는 본 발명의 일 실시예에 따른 초입방체 그라디언트 스케일 테이블의 일례를 설명하기 위한 도면이다.

[0061] 보다 자세히 설명하면, 초입방체 형상 테이블의 픽셀 위치 항을 n개의 파라미터별로 열로 나누어 나타내어 초입방체 그레이 스케일 테이블을 획득한다. 초입방체 그레이 스케일 테이블에서 특정 픽셀 A에 인접한 B(또는 C, D, E 등)와 그레이 스케일의 차이를 구하면, A 위치에서 A-B 선상의 그레이 스케일 변화율(기울기)을 구할 수 있다. 마찬가지로 나머지 방향으로 인접한 픽셀(C, D, E 등)들과의 그레이 스케일 차이를 구하면, 이 방향들(A-C, A-D, A-E 등)에 해당하는 그레이 스케일의 변화율(기울기)을 구할 수 있다. 이는 그라디언트를 의미하며, 각 픽셀에서 각 파라미터에 해당하는 그레이 스케일의 변화율(기울기)을 나타낸다. 이와 같은 방식으로 계산한 그라디언트 값을 초입방체 그레이 스케일 테이블에 추가 열로 붙여 초입방체 그라디언트 스케일 테이블을 획득한다.

[0062] 설명의 편의를 위해 도 11 및 도 12에 도시된 바와 같이, 각 파라미터의 좌표값은 이진수(0 또는 1)로 표시한다. 도 12의 맨 오른쪽 마지막 컬럼이 초입방체 그라디언트 스케일 테이블의 그라디언트 항목이다.

[0063] 도 12를 참조하면, 첫 번째 열의 첫 번째 행을 보면 "12-8=4"의 값으로 기울기를 구하고 있다. 그 값은 "Parameter_1"의 첫 번째 값 "8"과 두 번째 값 "12"의 "Gray_scale"의 값을 이용하여 기울기 값을 구한 것이다. 따라서, "12-8"의 값인 "4"가 된다. 즉, "Parameter_1"에 해당하는 그라디언트 값은, 해당하는 위치(첫 번째 열)와 "Parameter_1" 방향으로 인접한 위치, 즉 "Parameter"의 값이 "1(Parameter_1의 값), 0(Parameter_2의 값), 0(Parameter_3의 값), 0(Parameter_4의 값), 0(Parameter_5의 값), 0(Parameter_6의 값), 0(Parameter_7의 값)"인 위치의 그레이 스케일 값인 "12"에서, 해당하는 위치(첫 번째 열)의 그레이 스케일

일 값인 "8"을 빼서, 해당하는 위치(첫 번째 열)의 "Parameter_1"에 대한 그라디언트 값인 $12-8=4$ 를 구할 수 있다.

- [0064] 그라디언트 스케일의 두 번째 컬럼의 값은 "Parameter_2"의 첫 번째와 두 번째의 "Gray_scale" 값인 "16"과 "8"의 차인 "8"이다. 즉, "Parameter_2"에 해당하는 그라디언트 값은, 해당하는 위치(첫 번째 열)와 "Parameter_2" 방향으로 인접한 위치, 즉 "Parameter"의 값이 $0(\text{Parameter}_1\text{의 값}), 1(\text{Parameter}_2\text{의 값}), 0(\text{Parameter}_3\text{의 값}), 0(\text{Parameter}_4\text{의 값}), 0(\text{Parameter}_5\text{의 값}), 0(\text{Parameter}_6\text{의 값}), 0(\text{Parameter}_7\text{의 값})$ 인 위치의 그레이 스케일 값인 "16"에서, 해당하는 위치(첫 번째 열)의 그레이 스케일 값인 "8"을 빼서, 해당하는 위치(첫 번째 열)의 "Parameter_2"에 대한 그라디언트 값인 $16-8=8$ 를 구할 수 있다.
- [0065] 같은 방식으로 "Parameter_3"의 값은 $11-8=3$, "Parameter_4"의 값은 $9-8=1$, "Parameter_5"의 값은 $3-8=-5$, "Parameter_6"의 값은 $4-8=-4$, "Parameter_7"의 값은 $2-8=-6$ 으로 그라디언트 값을 구한다.
- [0066] 두번째 열, 세번째 열 등도 같은 방식으로 그라디언트를 구하여 전체 테이블을 만들게 된다.
- [0067] 군집 분석부(150)는 그라디언트 분석부(130)에 의해 획득된 초입방체 그라디언트 스케일 테이블을 기반으로 각 점에 대한 그라디언트 값을 이용하여 자동으로 군집을 분석한다.
- [0068] 군집 분석부(150)는 초입방체 그라디언트 스케일 테이블에서 그라디언트 항목의 모든 그라디언트 값이 0인 좌표 위치 항목을 획득할 수 있다. 그리고, 군집 분석부(150)는 획득한 좌표 위치 항목을 이용하여 자동으로 군집을 분석할 수 있다.
- [0069] 보다 자세히 설명하면, 군집 분석부(150)는 초입방체 그라디언트 스케일 테이블에서, 그레이 스케일 항목의 그레이 스케일 값이 0이 아니고 그라디언트 항목의 모든 그라디언트 값이 0인 좌표 위치는 군집의 대표값으로 선정하고, 그레이 스케일 항목의 그레이 스케일 값이 0이고 그라디언트 항목의 모든 그라디언트 값이 0인 좌표 위치는 군집의 경계값으로 선정할 수 있다. 그리고, 군집 분석부(150)는 군집의 대표값으로 선정된 좌표 위치와 군집의 경계값으로 선정된 좌표 위치를 이용하여 자동으로 군집을 분석할 수 있다.
- [0070] 도 13은 FCS 데이터를 히스토그램으로 나타낸 예시적인 도면이고, 도 14는 FCS 데이터를 입방체로 나타낸 예시적인 도면이다.
- [0071] 도 13을 참조하면 시각인식 기계학습과 별도로 상기 기법에 따라 FCS 데이터를 초입방체 형상으로 전환하면 멀티파라미터 FCS 데이터의 군집분석 시 수많은 히스토그램이나 점평면을 그리지 않고 단일 형상으로부터 직접 군집을 구별할 수 있다. 도 13에서 CD117, CD3/cCD3, CD19 히스토그램에서는 연속적인 게이팅(gating)을 통해 CD19+ 림프구(B 세포)를 선택하여 분석할 수 있다.
- [0072] 도 14를 참조하면, 같은 FCS 데이터를 입방체(또는 초입방체)로 나타낼 경우 한 번에 B 세포를 구분하여 선택할 수가 있다.
- [0073] 실제로 4차원 이상의 초입방체 형상을 시각적으로 표현해 군집을 확인하는 것은 불가능하기 때문에 초입방체 형상을 표현하는 데이터 테이블에서 같은 작업을 구현해야 한다. 이를 위해 초입방체 복셀의 각 좌표위치에 그레이 스케일 밀도 기울기 값을 부여하고 기울기 값이 0이 되는 부분을 군집의 최대밀도점 및 경계로 설정하여 군집을 설정할 수 있다.
- [0074] 도 15는 본 발명의 일 실시예에 따른 자동 군집 분석 과정을 설명하기 위한 도면이고, 도 16은 본 발명의 일 실시예에 따른 자동 군집 분석 과정에 이용되는 거리 계산 방법의 일례를 설명하기 위한 도면이다.
- [0075] 도 15와 같이 2-파라미터 데이터의 예를 통해 이를 나타내면, CD3와 CD19 마커로 T 림프구와 B 림프구를 정량분석한 예이다. 림프구 내 CD3+/CD19-, CD3-/CD19-, CD3-/CD19+ 세 군집을 확인할 수 있으며(도 15의 (a)), 군집 내 데이터포인트 밀도는 도 15의 (b)와 같은 양상을 띤다. 각 픽셀 좌표마다 그레이스케일 밀도 기울기를 배정하면 도 15의 (c)와 같이 군집 내 최대밀도점 주위의 기울기 값이 (0,0) 또는 (0,0)에 근사한 값이 된다. 이로부터 군집의 위치를 결정할 수 있다.
- [0076] 초입방체 그라디언트 스케일 테이블을 구하게 되면, 모든 기울기의 값이 0인 열이 나타나게 되는데, 이때 이 열의 값이 군집의 가장 중심에 있는 대표값과 가장 주변부의 경계값을 나타내는 것을 알 수 있다.
- [0077] 대표값과 주변부의 값을 구분하기 위해, "Gray_scale" 값이 0이 아닌 값으로 "Gradient_scale" 값이 0인 경우는 대표값으로, "Gray_scale" 값이 모두 0인 값으로 "Gradient_scale" 값이 0인 경우는 경계값으로 구분한다.

- [0078] 특정 군집의 대표값과 경계값을 구하게 되면, 그 군집에 속하는 좌표값을 찾기가 훨씬 용이하게 된다. 특정 군집의 대표값에서 경계값까지의 거리를 구한 다음에, 그 거리 안에 들어오는 모든 좌표값들은 같은 군집으로 분류될 수 있다. 용이한 시각적 표현과 발명의 이해를 위해 2-파라미터 데이터의 군집 분석의 예시는 도 15에 도시된 바와 같다. 도 15의 위쪽 방향 축은 아래쪽 평면에 dot plot으로 표시한 각 입자들의 수 및 그 분포를 나타낸다. 도 15의 아래쪽 그림은 그라디언트 0인 픽셀들의 위치로 파악한 각 군집의 대표 위치 및 범위를 나타낸다.
- [0079] 이때, 두 위치 사이의 거리는 도 16에 도시된 바와 같은 유클리드 거리를 이용하여 계산될 수 있다. 물론, 대표적인 비계층적인 군집화 알고리즘 중의 하나인 K-Means 방법을 이용하여 자동으로 군집을 찾을 수도 있다.
- [0080] 다시 설명하면, 분석 입자의 수가 충분히 많으면 n 차(n 개의 파라미터) 초입방체 공간 내에 표시된 입자들은 통계학적 분포를 보이며 군집을 형성하게 된다. 이에 따라 각 군집의 대표값 또는 대표 위치에 가장 많은 입자들이 위치하고 이 위치에서 멀어질수록 적은 수의 입자들이 분포할 것이다. 따라서, 초입방체 내 형상은 대표 위치에 해당하는 픽셀이 가장 높은 그레이 스케일 강도를 보이고, 주변부로 갈수록 강도가 점점 떨어지는 모양을 보인다. 평균 0인 일차원 정규 분포를 따르는 데이터의 경우 측정값 0에 해당하는 샘플의 수가 가장 많고 0에서 멀어질수록 분산에 따라 규칙적으로 샘플 수가 감소된다. N -파라미터 측정 자료도 마찬가지로 생각할 수 있다. 대표값에 해당하는 샘플의 수가 가장 많이 분포하고 각 파라미터 별로 대표값에서 멀어질수록 분포 특성에 의한 규칙에 따라 샘플의 수가 감소하는 양상을 보인다. 역으로 이 형상으로부터 각 군집의 대표 위치와 각 군집의 범위 또는 경계를 구하기 위해 통계적 밀도 분포 특성을 활용할 수 있다. 대표 위치에서 그레이 스케일 강도가 가장 높고 주변부로 갈수록 감소하는 특성 때문에 대표 위치에서 강도의 그라디언트 값은 0이 될 것이다. 또한, 군집 실제적 경계 밖에서는 강도의 변화가 없을 것이기 때문에 각 군집의 공간적 범위에 해당되지 않은 위치에서의 강도 그라디언트 또한 0의 값을 갖을 것이다. 그러므로, 초입방체 내 각 픽셀들의 강도 그라디언트를 구하고 이 중 0의 값을 보이는 픽셀들로부터 각 군집들의 대표 위치를 선별하고 그라디언트 0이 연속되는 구간으로부터 각 군집들의 범위를 선정할 수 있다.
- [0081] 질병 판단부(170)는 기계학습 알고리즘을 이용하여 군집 유형별 FCS 데이터를 통해 질병 판단 모델을 학습하고, 질병 판단 모델에 FCS 데이터를 입력하여 입력된 FCS 데이터에 대응되는 군집 유형을 획득할 수 있다.
- [0082] 보다 자세히 설명하면, FCS 데이터를 활용하여 기계학습용 데이터를 만들어서 인공지능을 활용한 질병 판단의 임상적 장치로 활용할 수 있다. 즉, 기계학습의 대표적인 모델로 군집화(Clustering)를 사용하여 입력 데이터들의 특징을 분석하여 서로 동일하거나 유사한 특징을 가진 데이터끼리 그룹화 한다. 따라서, 새로운 데이터가 입력되면 해당 데이터가 어느 군집에 속하는 지 분류가 가능하다. 예컨대, 혈액암 환자의 FCS 데이터를 군집화로 기계학습을 해놓으면, 이후에 특정한 FCS 데이터가 입력되었을 때 기존에 학습된 혈액암과 동일한 군집화의 모양을 보이게 되었을 때 혈액암으로 판단할 수 있다.
- [0083] 또한, 비지도학습의 군집화 방법을 사용하여 입력 데이터의 공간 군집 분석을 통해 타 질병과의 중첩성, 그리고 타 질병의 전이 가능성 등을 예측하는 모델로써도 사용할 수 있다.
- [0084] 도 17은 본 발명의 일 실시예에 따른 복셀 상태의 일례를 설명하기 위한 도면이다.
- [0085] 이러한 군집분석은, 초입방체 형상의 윤곽과 각 복셀의 그레이스케일 밀도를 결정하는 FCS 데이터 포인트들이 두 가지 조건을 만족해야 적용할 수 있다. 첫 번째 조건은 군집 내 데이터 포인트들이 확률적 분포를 하고 있으며 군집 내부 방향으로 데이터 포인트 밀도가 높아지는 것이고, 두 번째 조건은 데이터 포인트의 수가 확률 밀도를 충분히 반영할 정도로 충분히 많아야 한다. 하지만 이에 못 미치는 수의 데이터 포인트만 수집한 채 분석을 마쳐야 하는 경우가 빈번하다.
- [0086] 도 17의 (a)는 입방체 내에 충분한 수의 데이터 포인트들이 존재하기 때문에 복셀 간 농도 기울기 계산이 가능한 경우를 나타내며, 도 17의 (b)는 데이터 수집이 충분하지 않아 데이터 포인트로 점유된 복셀이 서로 떨어져 있는 상태를 나타낸다.
- [0087] 도 17의 (b)는 모든 복셀의 기울기가 ± 1 또는 0이고 군집의 밀도나 경계와 관련이 없어져 0-gradient를 이용한 군집분석이 불가능하다. 이를 극복하고 0-gradient 군집 분석을 적용하기 위해 초입방체 압축과 가상 데이터 포인트를 추가하는 방식을 적용한다.
- [0088] 도 17의 (b)와 같이 데이터 포인트로 점유된 유효 복셀(non-void voxel)이 서로 떨어져(isolate) 있는 경우, 복수의 복셀들을 병합(merge)함으로써 군집 내 비유효 복셀(void voxel)을 제거하고 농도 기울기(density

gradient)를 만들어 줄 수 있다. 이를 위해 초입방체 형상 테이블의 행을 구성하고 있는 복셀들의 좌표를 압축 배수에 해당하는 2 이상의 자연수로 나누고 그 몫으로 원 좌표를 대치한다.

- [0089] 도 18 및 도 19는 본 발명의 일 실시예에 따른 압축된 초입방체 변환 데이터의 일례를 설명하기 위한 도면이다. 도 18은 6-파라미터 FCS 데이터의 초입방체 형상 전환 테이블에서 초입방체 형상의 1/4배 압축이 어떻게 구현되는 지 나타낸 예이다.
- [0090] 첫 단계로 초입방체 복셀에 해당하는 각 행에 표시되어 있는 좌표 수치를 압축 배수인 4로 나눈다.
- [0091] 다음 단계로 도 19와 같이 새로 교체해준 몫의 소수점 아래자리 수를 버린다. 이러한 처리의 결과로 4 개씩의 행이 같은 좌표를 갖는 테이블이 생성된다. 그리고 같은 좌표를 갖는 4 개씩의 행들을 병합한다. 변환된 초입방체 형상 테이블의 좌표는 0 - 31의 수치를 갖는다. 이는 압축된 초입방체의 크기가 32^6 임을 나타낸다. 결과적으로 초입방체의 크기가 128^6 에서 32^6 로 줄어들었고 각 복셀 그레이스케일 밀도는 증가한다.
- [0092] 도 20은 본 발명의 일 실시예에 따른 압축된 초입방체 형상 데이터의 일례를 설명하기 위한 도면이다.
- [0093] 도 20은 같은 방법을 통해 2-파라미터 FCS 데이터에 해당하는 전환 이미지를 2배 압축할 때 나타나는 변화를 보여준다.
- [0094] 압축 결과, 원 초입방체 형상에 비해 압축 배수만큼 작지만 유사한 형태를 갖는 새로운 초입방체 형상이 생성된다. 압축 초입방체 형상은 군집 내부가 유효한(non-void) 복셀들로 촘촘해 채워지고 밀도가 높아 0-gradient 군집분석이 가능해진다. 다음 단계로 압축 초입방체 형상을 대상으로 0-gradient 군집분석을 수행하고 각 군집 간 최소 거리(D)를 구한다. 두 군집 간 거리는 각 군집에 속하는 복셀들 간 거리의 최소값으로 정의한다.
- [0095] 도 21 내지 도 26을 참조하여, 가상 데이터 포인트 추가 동작을 설명한다.
- [0096] 도 21은 본 발명의 일 실시예에 따른 초구체 설정 동작을 설명하기 위한 도면이다. 도 22는 본 발명의 일 실시예에 따른 벡터차의 방향을 설명하기 위한 도면이다. 도 23은 본 발명의 일 실시예에 따른 포인트 데이터 추가 동작을 설명하기 위한 도면이다. 도 24는 본 발명의 일 실시예에 따른 포인트 데이터 추가로 인한 그라디언트 기법 적용시 군집 상태 변화를 설명하기 위한 도면이다. 도 25는 본 발명의 일 실시예에 따른 초구체 직경과 군집 상태의 관계를 설명하기 위한 도면이다. 도 26은 본 발명의 일 실시예에 따른 최소 군집 간 거리를 초구체 직경으로 사용한 결과를 설명하기 위한 도면이다.
- [0097] 원 초입방체 형상에서 서로 가까이 위치하는 군집은 압축 과정에서 하나의 군집으로 서로 merge될 가능성이 있다. 그러므로 두 번째 단계에서는 가상 데이터 포인트를 원 형상에 추가하여 줌으로써 원 형상에서 비유효 복셀(void voxel)을 제거하고 군집의 밀도를 증가시킨다. 데이터 포인트를 추가한 원 규격의 초입방체에 0-gradient 군집분석을 시행하여 압축 형상에서 놓친 군집을 추가로 식별(identification)한다. 이를 위해 각 유효 복셀을 중심으로 하는 직경 D 또는 압축 배수의 초구체를 설정한다(도 21 참조).
- [0098] 이 초구체 범위 내에 존재하는 유효 복셀과 초구체 중심에 위치한 유효 복셀의 좌표로부터 각 복셀 쌍의 벡터차를 구한다(도 22 참조). 다음 각 초구체에서 구한 벡터차들을 합한다. 초구체 중심 유효 복셀이 군집의 변연에 위치할 경우 벡터차들은 도 22에 나타낸 것과 같이 일정한 방향성을 갖는다. 그러므로 벡터차들의 합은 군집 내부 방향으로 큰 절대값을 갖게 된다.
- [0099] 반면 초구체 중심 유효 복셀이 군집 내부에 위치할 경우 벡터차들은 다양한 방향성을 갖는다(도 22). 그러므로 벡터차들의 합은 방향이 일정하지 않으면서 매우 작은 절대값을 갖게 된다.
- [0100] 다음 단계에 초구체 중심으로부터 벡터차 합의 방향과 거리에 위치한 복셀에 데이터 포인트를 추가한다(도 23 참조). 초구체 중심이 군집 변연에 있는 유효 복셀일 경우 군집 내부로 향하는 위치에 데이터 포인트가 추가된다. 물론 초구체 중심이 군집 내부에 있는 경우도 벡터차 합의 절대값이 작기 때문에 추가되는 데이터 포인트는 군집 내부에 위치하게 된다. 이렇게 모든 원 초입방체 형상의 유효 복셀들에 대해 같은 작업을 하면 처음 존재하던 유효 복셀의 2 배에 해당하는 데이터 포인트들이 추가 유효 복셀로 더해지게 된다.
- [0101] 이러한 작업 전체를 n 번 반복하게 되면 주로 군집 중심 쪽에 모두 2^{n-1} 개의 유효 복셀들이 초입방체에 더해지게 되고 초입방체 형상은 음영이 변화를 갖는 부드러운 모양을 띠어 0-gradient 기법으로 군집의 식별(identification)이 가능해지게 된다(도 24 참조).
- [0102] 초구체의 직경으로 압축 초입방체 군집 간 최소 거리(D) 또는 압축배수를 사용한다. 초구체의 직경이 너무 클

경우 두 가지 문제가 발생한다. 우선 벡터차를 구해야 하는 유효 복셀 쌍의 수가 너무 많아져 연산 부담이 커질 수 있다. 또한 군집 외곽의 유효 복셀을 중심으로 하는 초구체에 인접한 다른 군집의 유효 복셀이 포함될 경우, 도 25와 같이 다른 군집 복셀과의 벡터차가 절대값이 크고 방향이 소속 군집과 반대 방향이기 때문에 벡터차 합의 말단이 두 군집 사이에 위치하게 된다. 그러므로, 추가되는 데이터 포인트가 두 군집 사이에 위치하게 되기 때문에 군집 간 구분을 모호하게 할 수 있다.

[0103] 압축 초입방체 군집분석에서 구한 최소 군집 간 거리 D를 초구체 직경으로 사용하는 것은, 압축 초입방체 군집 분석으로 식별한 군집들을 원 초입방체 군집분석 과정에서 계속 보존하기 위한 목적을 갖고 있다. 압축 배수를 초구체 직경으로 사용하는 것은 압축 과정에서 사라질 수 있는 군집 간 구분 원 초입방체 군집분석에서 되살리기 위해서이다. 도 26에서 16 거리만큼 떨어져 있는 두 군집을 육안으로 확인할 수 있으나 1/16 배 압축 결과 구분이 되지 않는 문제가 있다.

[0104] 본 실시예에 따른 데이터 처리 장치는 초구체의 직경을 압축된 초입방체 군집 간 최소 거리 또는 압축 배수로 설정하여 이러한 문제들을 해결한다.

[0105] 도 27은 본 발명의 일 실시예에 따른 질병 판단 모델의 기계학습에 이용되는 학습 데이터의 일례를 설명하기 위한 도면이다.

[0106] 이때, 질병 판단부(170)는 초입방체 형상 테이블에 적절한 필터를 처리하여 기계학습용 데이터로 전환할 수도 있다.

[0107] 예컨대, 질병 판단부(170)는 도 27에 도시된 바와 같이, 초입방체 형상 테이블의 데이터를 2차원적 재배열하여 이미지 유사 형식으로 전환하여 종래의 CNN 기계학습을 전이학습 형태로 직접 적용할 수도 있다. 즉, 질병 판단부(170)는 초입방체 형상 테이블에 포함된 좌표값과 카운트값을 병합하고, 병합된 값으로 1차원 좌표값을 구성한다. 이후, 질병 판단부(170)는 1차원 좌표값을 기반으로 데이터를 재구성하여 기계학습용 데이터 테이블을 생성한다. 예를 들어, 질병 판단부(170)는 1차원 좌표값을 2차원 배열로 재구성할 수 있다. 질병 판단부(170)는 변형된 데이터에 포함된 좌표값을 1차원 좌표값으로 구성하고, 1차원 좌표값으로 구성하는 과정에서 좌표값이 존재하지 않는 부분을 0 값으로 채우거나, 또는 좌표값이 존재하는 부분만 표시하는 방법을 사용하여 $\prod_{i=1}^m n_i$ (n_i 는 기 설정된 기준 크기값 이상의 자연수) 형태의 기계학습용 데이터 테이블로 재구성할 수 있다.

[0108] 도 28 내지 도 31은 본 발명의 다른 실시예에 따른 데이터 처리 방법을 예시한 흐름도이다. 데이터 처리 방법은 데이터 처리 장치에 의해 수행될 수 있다.

[0109] 데이터 처리 방법은 초기 데이터를 초입방체 공간 내 좌표와 농도 데이터를 갖는 초입방체 형상 데이터로 변환하고, 초입방체 형상 데이터를 테이블 기반의 초입방체 변환 데이터로 표현하는 데이터 변형 단계(S10), 초입방체 형상 데이터 또는 초입방체 변환 데이터를 기반으로 분석을 수행하여 분석 결과를 출력하는 분석 단계(S20)를 포함한다.

[0110] 초기 데이터는 혈액의 유세포 분석(Flow Cytometry)에 대한 결과를 나타내는 생체 추출 데이터를 포함하며, 생체 추출 데이터는 미리 설정된 정형화된 형식 또는 FCS(Flow Cytometry Standard) 형식으로 표현된다. 초입방체 변환 데이터는 복수의 축에 해당하는 좌표 정보 및 좌표 정보에 매칭하는 값 정보가 행 방향 또는 열 방향을 기준으로 배치되는 테이블로 표현된다.

[0111] 데이터 변형 단계(S10)는, 초입방체 형상 데이터 또는 초입방체 변환 데이터를 압축 배수를 기준으로 압축하는 단계(S11)를 포함할 수 있다.

[0112] 압축하는 단계(S11)는 초입방체 변환 데이터에 대해서 압축 배수로 나눈 몫을 기준으로 복수의 라인 병합(merge)을 통해 농도를 증가시킬 수 있다. 분석 단계(S20)는 압축된 초입방체 형상 데이터 또는 압축된 초입방체 변환 데이터에 대해서 그라디언트 군집 분석을 수행할 수 있다.

[0113] 데이터 변형 단계(S10)는 초입방체 형상 데이터 또는 초입방체 변환 데이터에 가상 데이터 포인트를 추가하는 단계(S12)를 포함할 수 있다.

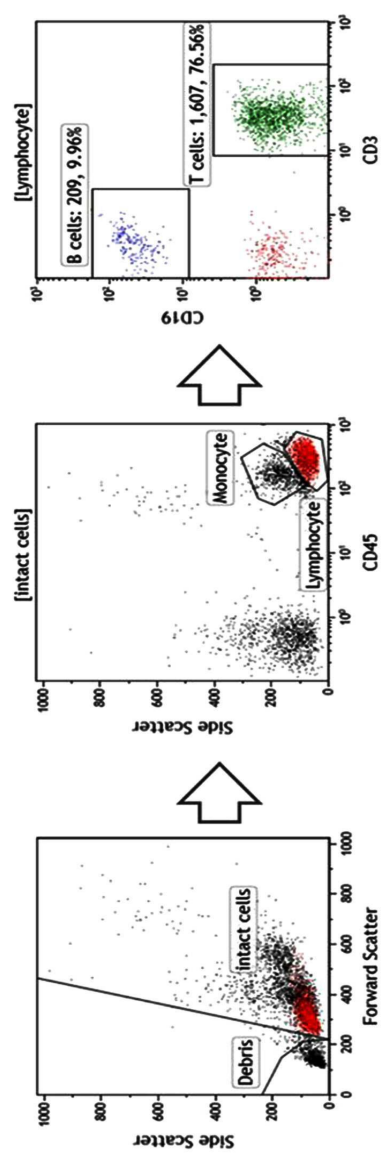
[0114] 가상 데이터 포인트를 추가하는 단계(S12)는, 유효 복셀을 중심으로 하는 초구체를 설정하고(S121), 초구체의 범위 내에 존재하는 유효 복셀과 상기 초구체의 중심에 위치한 유효 복셀의 벡터차를 산출하고(S122), 각 초구체에서 산출한 벡터차를 합하고(S123), 초구체의 중심으로부터 벡터차의 합의 방향과 거리에 위치한 복셀에 포인트 데이터를 추가하는(S124) 단계를 수행할 수 있다.

- [0115] 가상 데이터 포인터를 추가하는 단계(S12)는, 군집 내부로 향하는 위치 또는 군집 내부에 유효 복셀을 추가하는 과정을 $N(N \text{은 } 2 \text{이상 자연수})$ 번 반복하여 2^{n-1} 개의 유효 복셀을 추가할 수 있다.
- [0116] 가상 데이터 포인터를 추가하는 단계(S12)는, 초구체의 직경을 압축된 초입방체 군집 간 최소 거리 또는 압축 배수로 설정할 수 있다.
- [0117] 도 31을 참조하여 본 발명의 일 실시예에 따른 FCS 데이터의 초입방체 형상 테이블을 기초로 초입방체 픽셀의 그라디언트 분석을 통해 자동으로 군집을 분석하는 방법에 대하여 설명한다.
- [0118] 도 31은 본 발명의 일 실시예에 따른 FCS 데이터의 초입방체 형상 테이블을 기초로 초입방체 픽셀의 그라디언트 분석을 통해 자동으로 군집을 분석하는 방법을 설명하기 위한 흐름도이다.
- [0119] 도 31를 참조하면, 자동 군집 분석 장치(100)는 FCS 데이터를 기반으로 각 입자의 파라미터 측정값들을 이용하여 각 입자를 각 파라미터를 축으로 하고 파라미터의 측정값들을 좌표 위치로 하는 초입방체의 점으로 표현하는 초입방체 형상 테이블을 획득한다(S110).
- [0120] 즉, 자동 군집 분석 장치(100)는 FCS 데이터의 파라미터 측정값들을 각 입자별로 병합하여 각 입자의 좌표 위치를 획득할 수 있다. 이때, 자동 군집 분석 장치(100)는 FCS 데이터의 전체 파라미터 중에서 적어도 2개 이상의 파라미터에 대한 측정값들을 각 입자별로 병합하여 각 입자의 좌표 위치를 획득할 수 있다. 또한, 자동 군집 분석 장치(100)는 FCS 데이터에 포함된 전체 또는 일부 파라미터의 측정값들을 미리 선정된 일정한 값으로 나누어 몫으로 치환하는 방법으로 데이터를 변형(이미지 깊이 변환)하고, 이때 발생하는 데이터의 유실을 막기 위해 각 몫에 소정의 값을 더하는 방법을 취할 수 있다.
- [0121] 그리고, 자동 군집 분석 장치(100)는 획득한 각 입자의 좌표 위치를 크기순으로 재배열할 수 있다.
- [0122] 그리고, 자동 군집 분석 장치(100)는 좌표 위치에 대응되는 그레이 스케일 값을 해당 좌표 위치를 가지는 입자의 개수를 기초로 획득하여, 좌표 위치 항목과 그레이 스케일 항목으로 이루어지는 초입방체 형상 테이블을 획득할 수 있다.
- [0123] 그런 다음, 자동 군집 분석 장치(100)는 초입방체 형상 테이블을 기반으로 각 점의 좌표 위치를 파라미터별로 분리하여 표현하는 초입방체 그레이 스케일 테이블을 획득하고, 초입방체 그레이 스케일 테이블을 기반으로 각 점별로 그라디언트를 분석하여 각 점별 그라디언트 값을 포함하는 초입방체 그라디언트 스케일 테이블을 획득한다(S130).
- [0124] 자동 군집 분석 장치(100)는 초입방체 형상 테이블의 좌표 위치 항목을 파라미터별로 분리하여 초입방체 그레이 스케일 테이블을 획득할 수 있다.
- [0125] 자동 군집 분석 장치(100)는 획득된 초입방체 그레이 스케일 테이블의 각 좌표 위치 항목에 대해 파라미터별로 그라디언트를 분석하여 파라미터별 그라디언트 값을 획득하여, 복수개의 파라미터별로 분리되어 있는 좌표 위치 항목, 그레이 스케일 항목 및 그라디언트 항목으로 이루어지는 초입방체 그라디언트 스케일 테이블을 획득할 수 있다.
- [0126] 자동 군집 분석 장치(100)는 초입방체 그레이 스케일 테이블의 각 좌표 위치 항목에 대해, 해당 좌표 위치 항목에 대응되는 그레이 스케일 값과 파라미터별로 인접한 복수개의 좌표 위치 항목에 대응되는 복수개의 그레이 스케일 값을 이용하여 해당 좌표 위치 항목에 대한 파라미터별 그라디언트 값을 획득할 수 있다.
- [0127] 그런 다음, 자동 군집 분석 장치(100)는 초입방체 그라디언트 스케일 테이블을 기반으로 각 점에 대한 그라디언트 값을 이용하여 자동으로 군집을 분석한다(S150).
- [0128] 자동 군집 분석 장치(100)는 초입방체 그라디언트 스케일 테이블에서 그라디언트 항목의 모든 그라디언트 값이 0인 좌표 위치 항목을 획득할 수 있다. 그리고, 자동 군집 분석 장치(100)는 획득한 좌표 위치 항목을 이용하여 자동으로 군집을 분석할 수 있다.
- [0129] 보다 자세히 설명하면, 자동 군집 분석 장치(100)는 초입방체 그라디언트 스케일 테이블에서, 그레이 스케일 항목의 그레이 스케일 값이 0이 아니고 그라디언트 항목의 모든 그라디언트 값이 0인 좌표 위치는 군집의 대표값으로 선정하고, 그레이 스케일 항목의 그레이 스케일 값이 0이고 그라디언트 항목의 모든 그라디언트 값이 0인 좌표 위치는 군집의 경계값으로 선정할 수 있다. 그리고, 자동 군집 분석 장치(100)는 군집의 대표값으로 선정된 좌표 위치와 군집의 경계값으로 선정된 좌표 위치를 이용하여 자동으로 군집을 분석할 수 있다.

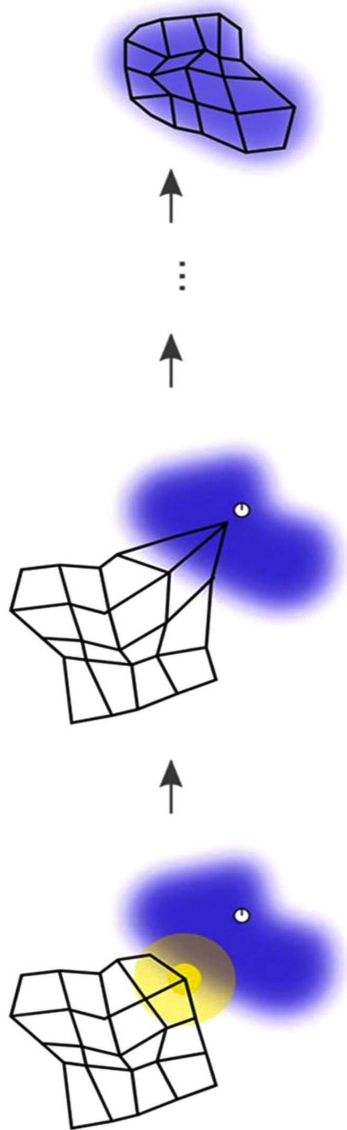
- [0130] 이후, 자동 군집 분석 장치(100)는 기계학습 알고리즘을 이용하여 군집 유형별 FCS 데이터를 통해 질병 판단 모델을 학습하고, 질병 판단 모델에 FCS 데이터를 입력하여 입력된 FCS 데이터에 대응되는 군집 유형을 획득할 수 있다(S170).
- [0131] 본 발명은 n-차원 FCS 데이터를 n-차원 초입방체 내 형상을 나타내는 테이블로 전환하여 그라디언트 분석을 통해 자동으로 군집을 분석함으로써, 생물학적/임상적 응용을 위해 이용될 수 있다. 생물학 및 의학 분야에서 FCS 데이터의 분석은 군집 분석을 의미한다. 이는 시간 순으로 배열된(입자 또는 세포 별) 데이터 개체의 파라미터 정보를 검토하여 유사한 파라미터 정보를 보이는 개체들을 모아 하나의 군집으로 분류하고 각 군집의 크기(개체 수)를 분석하는 작업이다. FCS 데이터를 생성하는 생물학/의료 분야의 분석/진단 장비는 이미 전 세계적으로 널리 보급되어 저차원의 기초적 분석이나 고차원의 심층 분석 검사에 통상적으로 사용되고 있다. 이러한 예로서 범용의 유세포 분석 장비와 자동 혈구 분석 장비가 있다. 유세포 데이터나 자동 혈구 분석 검사의 분석에 있어 본 발명은 다음과 같은 장점을 갖고 있다.
- [0132] 1. 유세포 분석 기법의 발전 방향이 파라미터 수의 증가로 보다 자세한 분류/분석을 가능하게 하는 데에 있고, 이에 따라 FCS 데이터 구조가 복잡해져 신속한 수기 분석과 적절한 형식의 결과 보고가 점차 힘들어지고 있다. 이러한 배경 하에, 멀티-파라미터 유세포 분석 데이터의 군집 분석을 자동으로 시행하는 SOM 기법에 기반한 알고리즘과 서비스가 소개/보급되고 있다. 이러한 분석 기법은 시각화를 거치지 않아 분석 원리 및 분석 결과를 이해하고 파악하는 것이 어렵다. 본 발명은 FCS 데이터 전체를 형상화한 후 이를 대상으로 분석을 진행하기 때문에, 고전적인 시각적 수기 군집 분석법과 유사성을 갖고 있다. 따라서, 분석자가 분석 원리나 분석 결과를 보다 쉽고 직관적으로 이해할 수 있다.
- [0133] 2. 자동 혈구 분석 검사는 제한된 시간 안에 많은 검체의 검사를 소화해야 하기 때문에 FCS 데이터의 군집 분석을 시행함에 있어 다른 유세포 분석 데이터의 분석과 다소 상이한 분석 기법을 사용한다. 자동 혈구 분석은 혈액에 존재하는 세포의 종류, 즉 군집의 수를 미리 결정된 상태로 군집 분석을 시행하고 각 군집의 크기(세포 수)를 보고한다. 이러한 분석 방식은 분석 효율이 높은 반면 혈구 세포들의 변이와 특이 세포군의 존재를 감지하지 못하는 단점이 있다. 본 발명은 FCS 데이터의 기계학습 인공지능 분석 적용을 가능하게 함과 동시에, 혈구 세포들의 종류와 그 변이, 그리고 특이 세포군의 존재를 민감하게 감지하고 보고하는 것을 가능하게 한다.
- [0134] 3. 본 발명은 유세포 기반 분석 장비들의 분석 데이터인 FCS 데이터를 향상된 방식으로 분석할 수 있게 해주는 것을 목적으로 하고 있다. 따라서, 전 세계 생물학 연구 시설과 의료 시설에 광범위하게 보급되어 사용되고 있는 유세포 분석 장비와 자동 혈구 분석 장비의 add-on 장치로서 쉽게 시장 접근이 가능하여 큰 사업 확장성을 갖을 것으로 예상된다.
- [0135] 이상에서 설명한 본 발명의 실시예를 구성하는 모든 구성요소들이 하나로 결합하거나 결합하여 동작하는 것으로 기재되어 있다고 해서, 본 발명이 반드시 이러한 실시예에 한정되는 것은 아니다. 즉, 본 발명의 목적 범위 안에서라면, 그 모든 구성요소들이 하나 이상으로 선택적으로 결합하여 동작할 수도 있다. 또한, 그 모든 구성요소들이 각각 하나의 독립적인 하드웨어로 구현될 수 있지만, 각 구성요소들의 그 일부 또는 전부가 선택적으로 조합되어 하나 또는 복수개의 하드웨어에서 조합된 일부 또는 전부의 기능을 수행하는 프로그램 모듈을 갖는 컴퓨터 프로그램으로서 구현될 수도 있다. 또한, 이와 같은 컴퓨터 프로그램은 USB 메모리, CD 디스크, 플래쉬 메모리 등과 같은 컴퓨터가 읽을 수 있는 기록 매체(Computer Readable Media)에 저장되어 컴퓨터에 의하여 읽혀지고 실행됨으로써, 본 발명의 실시예를 구현할 수 있다. 컴퓨터 프로그램의 기록 매체로서는 자기기록매체, 광기록매체 등이 포함될 수 있다.
- [0136] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위 내에서 다양한 수정, 변경 및 치환이 가능할 것이다. 따라서, 본 발명에 개시된 실시예 및 첨부된 도면들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예 및 첨부된 도면에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

도면

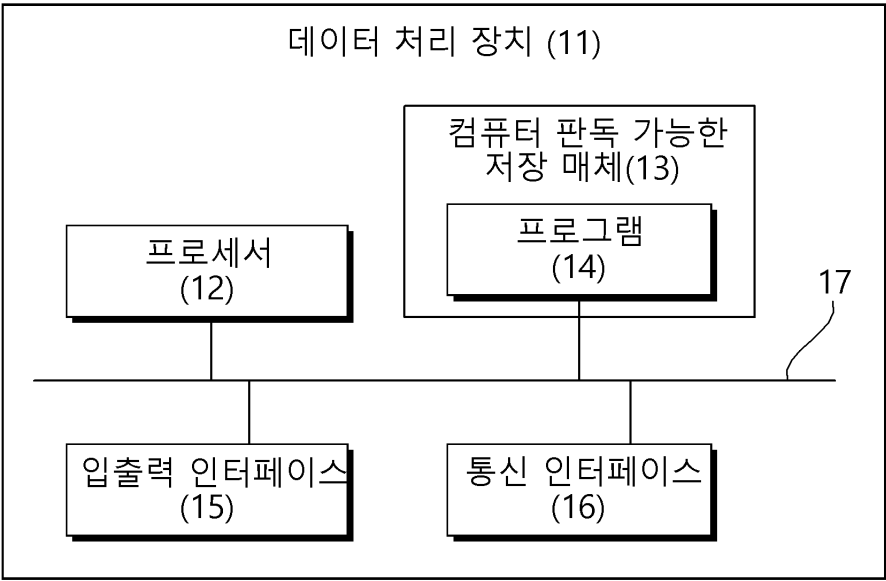
도면1



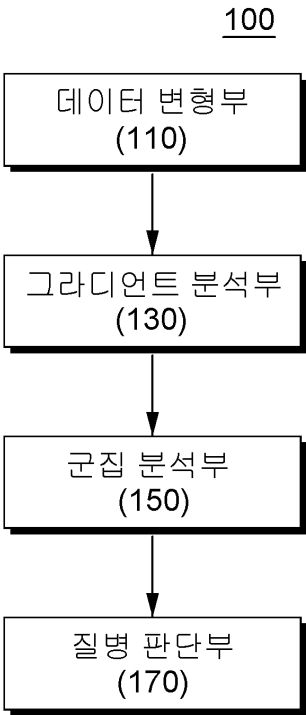
도면2



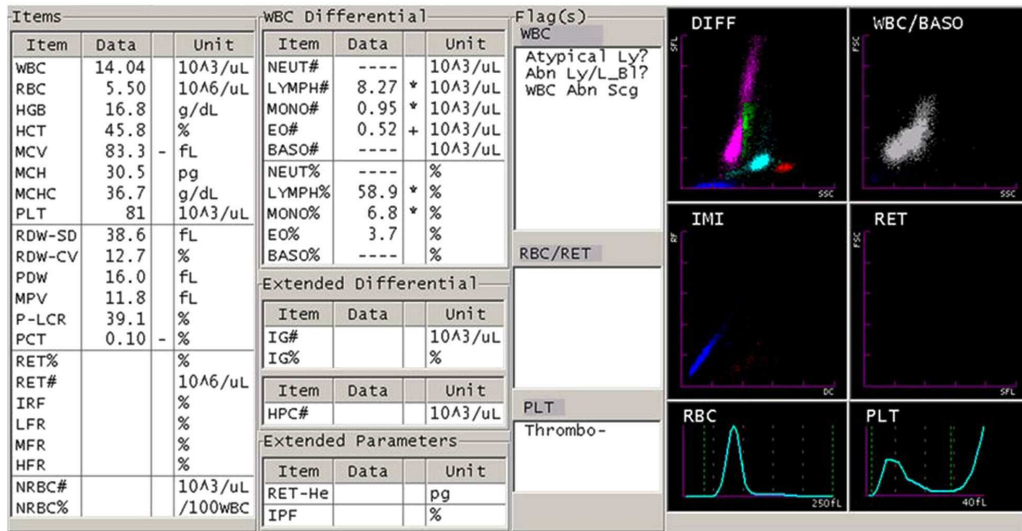
도면3



도면4



도면5



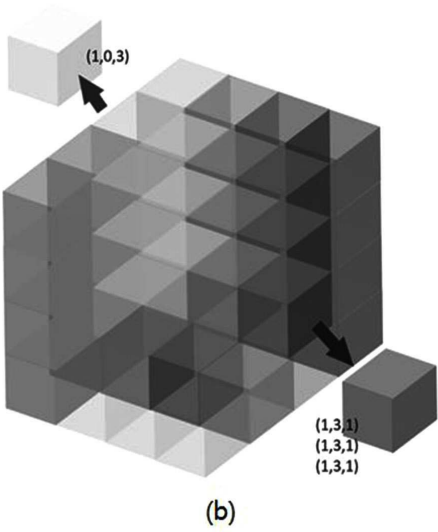
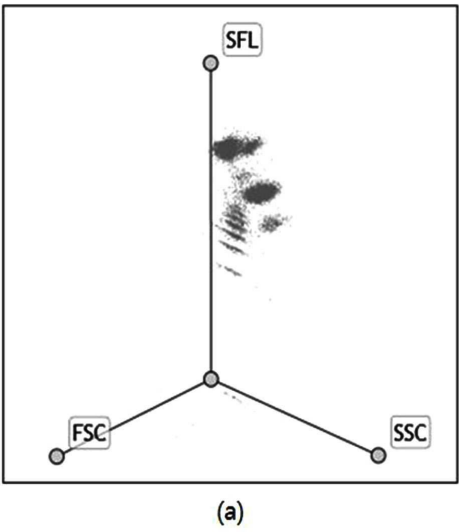
도면6

파일 이름: a.fcs

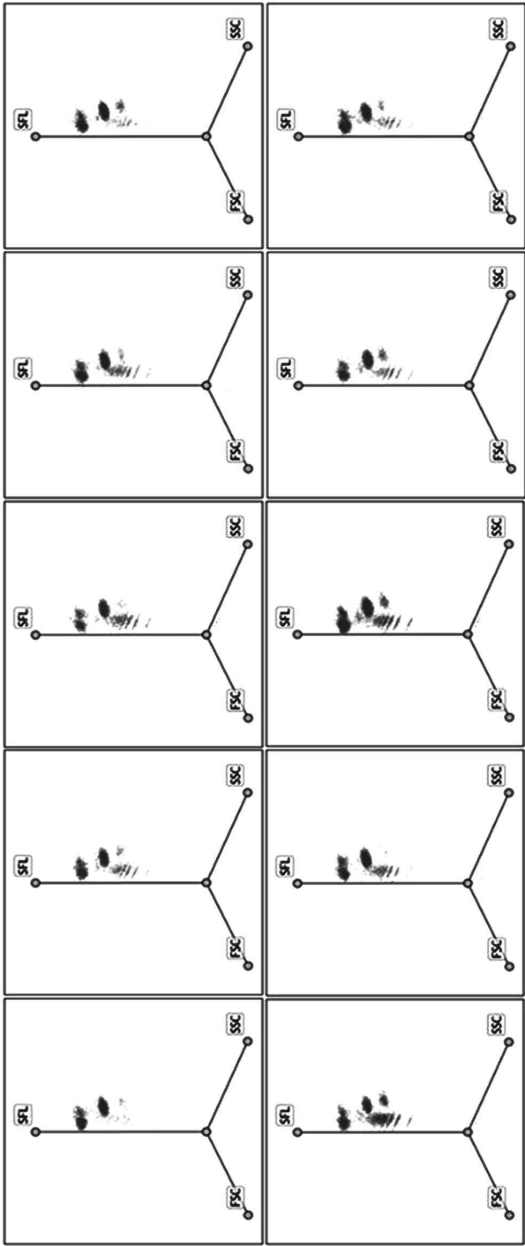
검사일련번호 : 111-222-333, 검사일: yy-mm-dd-hh

AAA	BBB	CCC	DDD
125	205	11	5
34	61	81	75
34	61	81	75
° °	° °	° °	° °

도면7



도면8



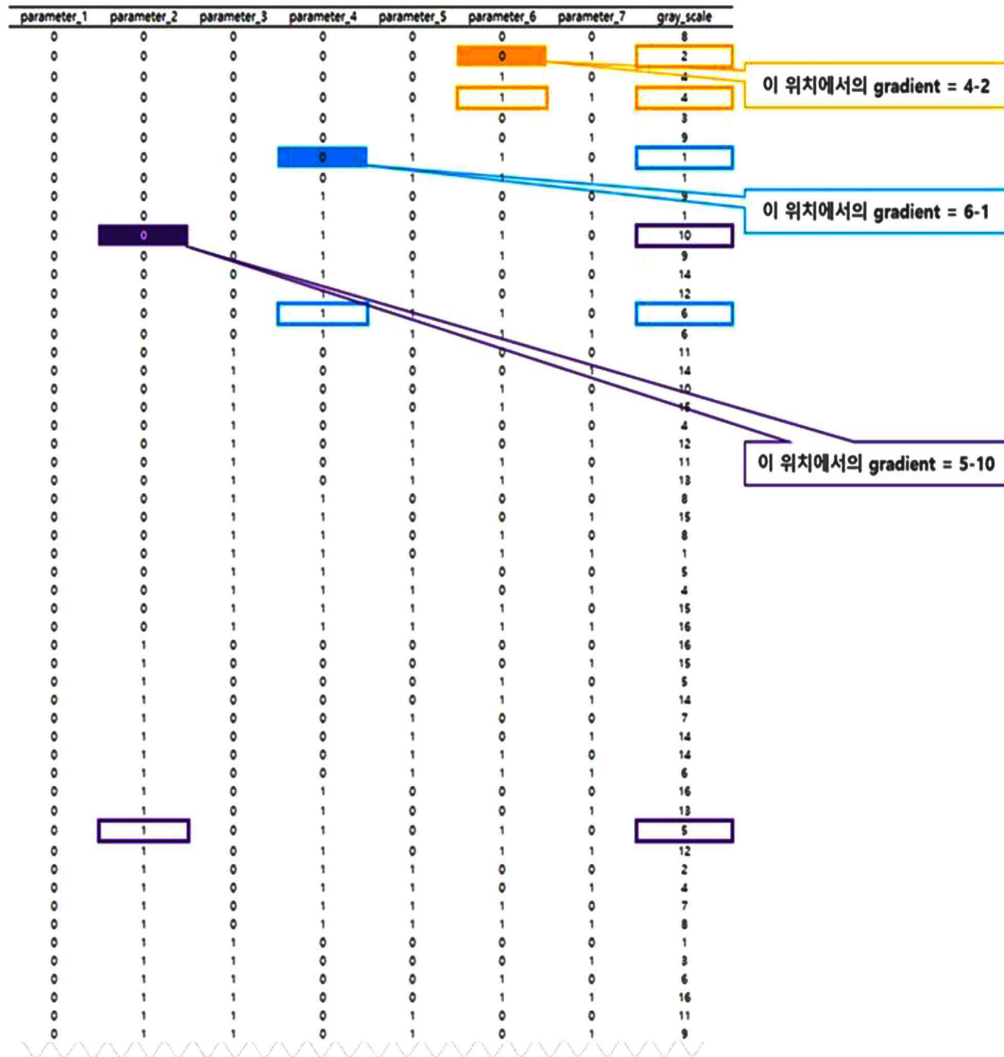
도면9

AAABBBCCDDDD	Count
125205115(좌표를 나타냄)	1
34618175	2

도면10

Par. 1	Par. 2	Par. 3	Par. 4	Par. 5	Par. 6	...	Par. n	Density
0	0	0	0	0	0		0	0
1	0	0	0	0	0		0	2
2	0	0	0	0	0		0	16
3	0	0	0	0	0		0	11
4	0	0	0	0	0		0	3
0	1	0	0	0	0		0	1
1	1	0	0	0	0		0	18
2	1	0	0	0	0		0	14
3	1	0	0	0	0		0	1
4	1	0	0	0	0		0	0
0	2	0	0	0	0		0	1
1	2	0	0	0	0		0	16
2	2	0	0	0	0		0	12
3	2	0	0	0	0		0	0
4	2	0	0	0	0		0	1

도면11

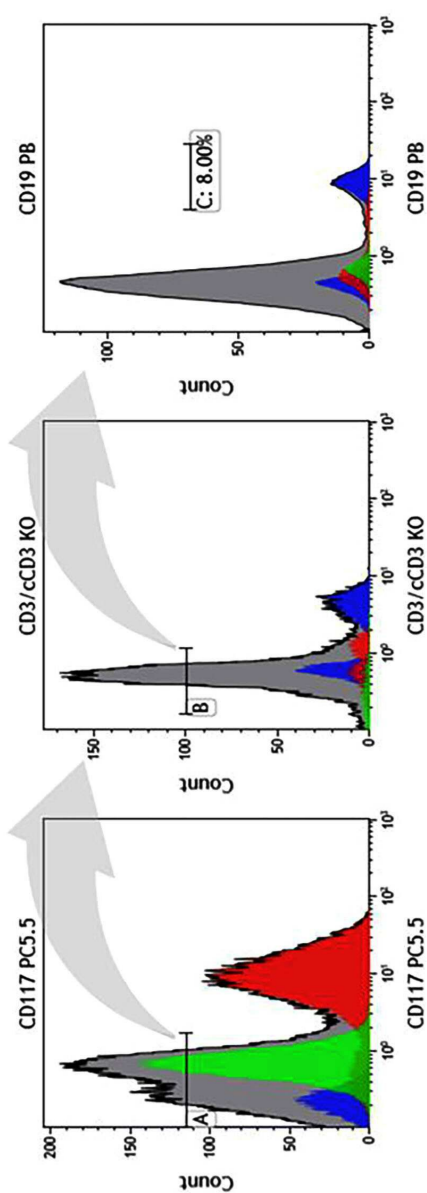


도면12

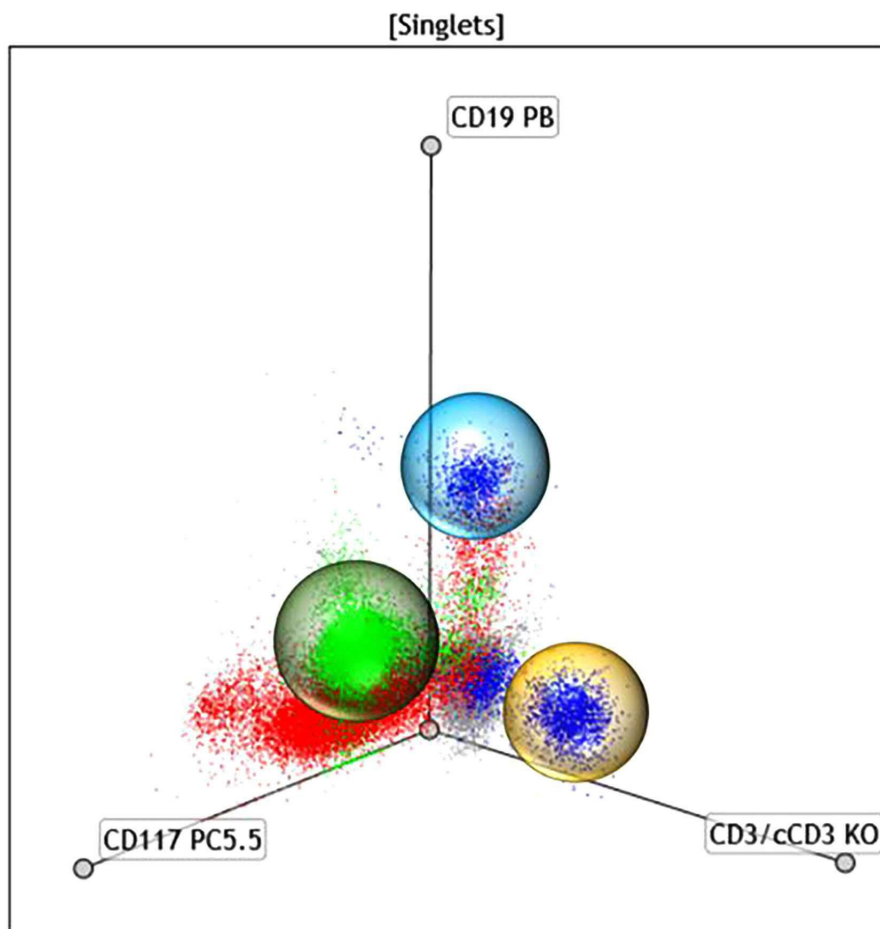
parameter_1	parameter_2	parameter_3	parameter_4	parameter_5	parameter_6	parameter_7	gray_scale	gray_scale_gradient
0	0	0	0	0	0	0	8	(4,8,3,1,-5,-4,-6)
0	0	0	0	0	0	1	2	
0	0	0	0	0	1	0	4	
0	0	0	0	0	1	1	4	
0	0	0	0	1	0	0	3	
0	0	0	0	1	0	1	9	
0	0	0	0	1	1	0	1	
0	0	0	1	1	1	1	1	
0	0	0	1	0	0	0	9	
0	0	0	1	0	0	1	1	
0	0	0	1	0	1	0	10	
0	0	0	1	0	1	1	9	
0	0	0	1	1	0	0	14	
0	0	0	1	1	0	1	12	
0	0	0	1	1	1	0	6	
0	0	0	1	1	1	1	6	
0	0	1	0	0	0	0	11	
0	0	1	0	0	0	1	14	
0	0	1	0	0	1	0	10	
0	0	1	0	0	1	1	15	
0	0	1	0	1	0	0	4	
0	0	1	0	1	0	1	12	
0	0	1	0	1	1	0	11	
0	0	1	0	1	1	1	13	
0	0	1	1	0	0	0	8	
0	0	1	1	0	0	1	15	
0	0	1	1	0	1	0	8	
0	0	1	1	0	1	1	1	
0	0	1	1	1	0	0	5	
0	0	1	1	1	0	1	4	
0	0	1	1	1	1	0	15	
0	0	1	1	1	1	1	16	
0	1	0	0	0	0	0	16	
0	1	0	0	0	0	1	15	
0	1	0	0	0	1	0	5	
0	1	0	0	0	1	1	14	
0	1	0	0	1	0	0	7	
0	1	0	0	1	0	1	14	
0	1	0	0	1	1	0	14	
0	1	0	0	1	1	1	6	
0	1	0	1	0	0	0	16	
0	1	0	1	0	0	1	13	
0	1	0	1	0	1	0	5	
0	1	0	1	0	1	1	12	
0	1	0	1	1	0	0	2	
0	1	0	1	1	0	1	4	
0	1	0	1	1	1	0	7	
0	1	0	1	1	1	1	8	
0	1	1	0	0	0	0	1	
0	1	1	0	0	0	1	3	
0	1	1	0	0	1	0	6	
0	1	1	0	0	1	1	16	
0	1	1	0	1	0	0	11	

(12-8, 16-8, 11-8, 9-8, 3-8, 4-8, 2-8)

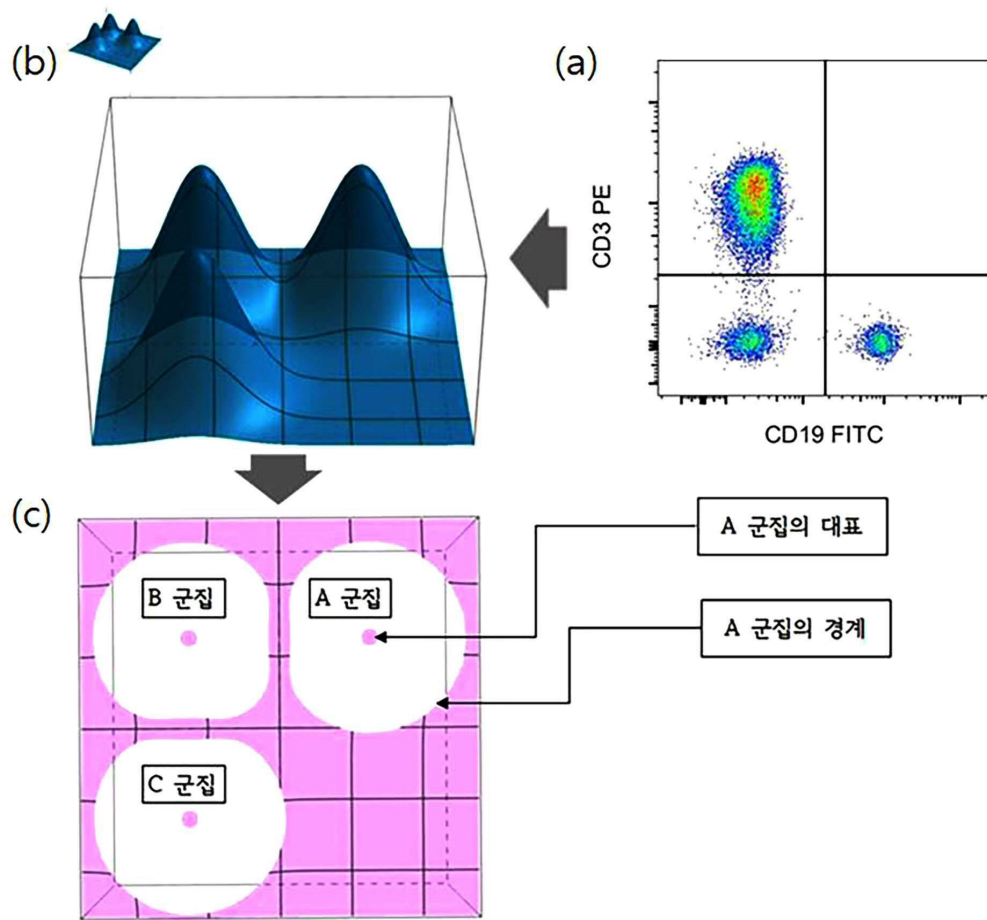
도면13



도면14



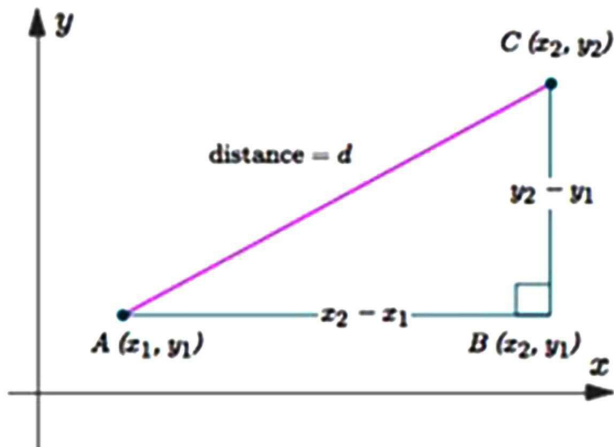
도면15



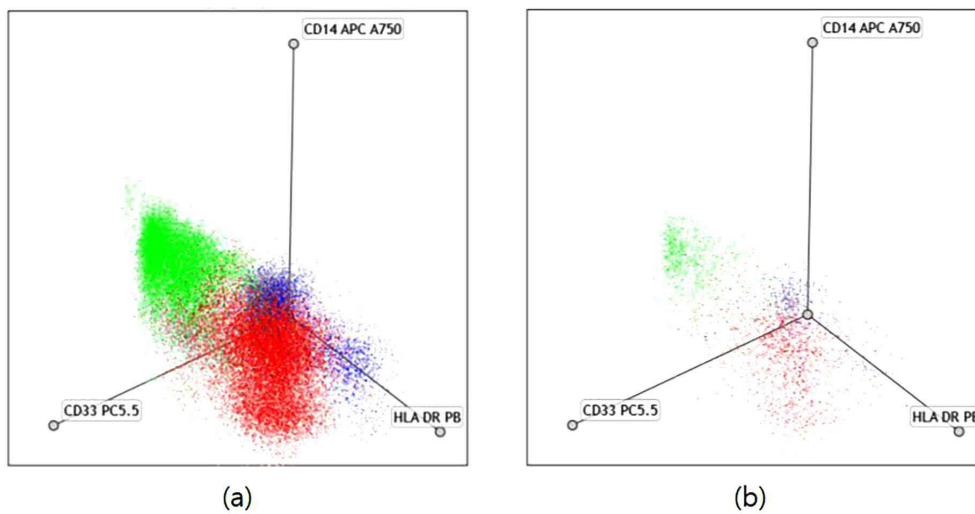
도면16

- 두 점 사이의 거리를 계산하는 방법
- 이 거리를 이용하여 유클리드 공간 정의

$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$



도면17



도면18

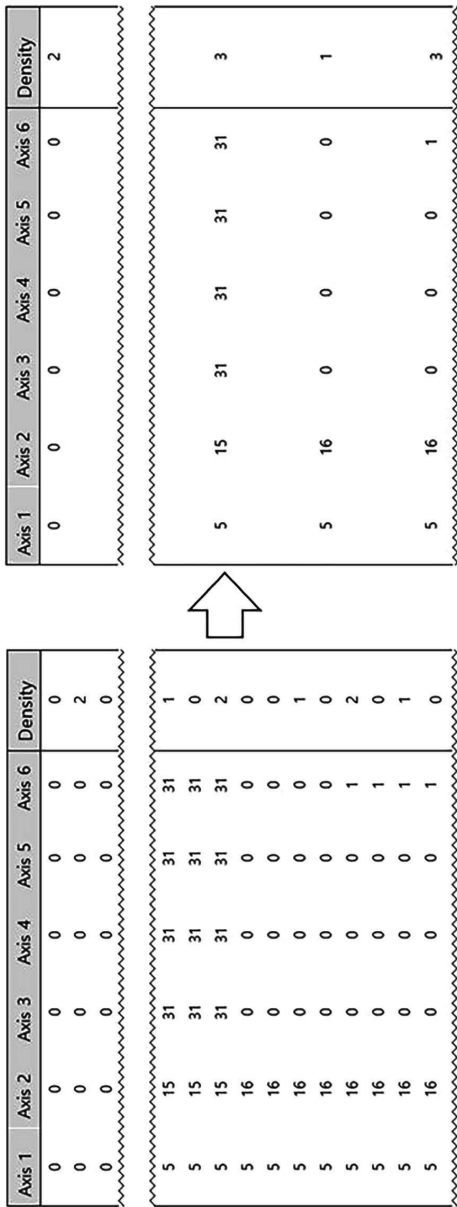
Axis 1	Axis 2	Axis 3	Axis 4	Axis 5	Axis 6	Density
0	0	0	0	0	0	0
0	0	0	0	0	1	2
0	0	0	0	0	2	0

21	63	127	127	127	125	1
21	63	127	127	127	126	0
21	63	127	127	127	127	2
21	64	0	0	0	0	0
21	64	0	0	0	1	0
21	64	0	0	0	2	1
21	64	0	0	0	3	0
21	64	0	0	0	4	2
21	64	0	0	0	5	0
21	64	0	0	0	6	1
21	64	0	0	0	7	0

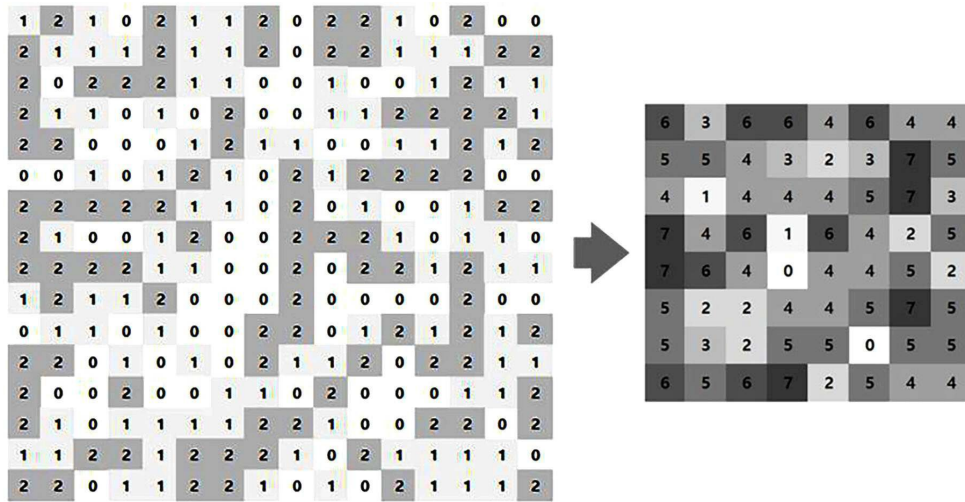
Axis 1	Axis 2	Axis 3	Axis 4	Axis 5	Axis 6	Density
0.00	0.00	0.00	0.00	0.00	0.00	0
0.00	0.00	0.00	0.00	0.00	0.25	2
0.00	0.00	0.00	0.00	0.00	0.50	0

525	15.75	31.75	31.75	31.75	31.25	1
525	15.75	31.75	31.75	31.75	31.50	0
525	15.75	31.75	31.75	31.75	31.75	2
525	16.00	0.00	0.00	0.00	0.00	0
525	16.00	0.00	0.00	0.00	0.25	0
525	16.00	0.00	0.00	0.00	0.50	1
525	16.00	0.00	0.00	0.00	0.75	0
525	16.00	0.00	0.00	0.00	1.00	2
525	16.00	0.00	0.00	0.00	1.25	0
525	16.00	0.00	0.00	0.00	1.50	1
525	16.00	0.00	0.00	0.00	1.75	0

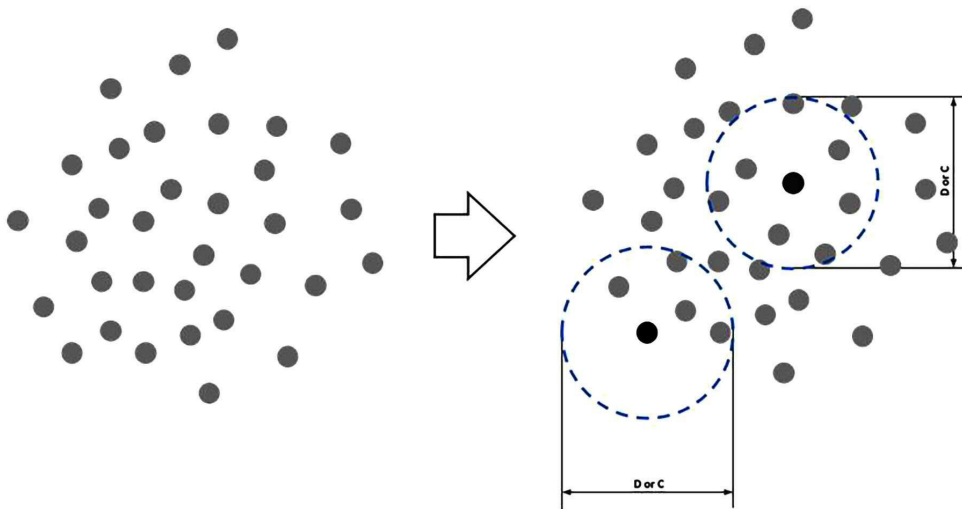
도면19



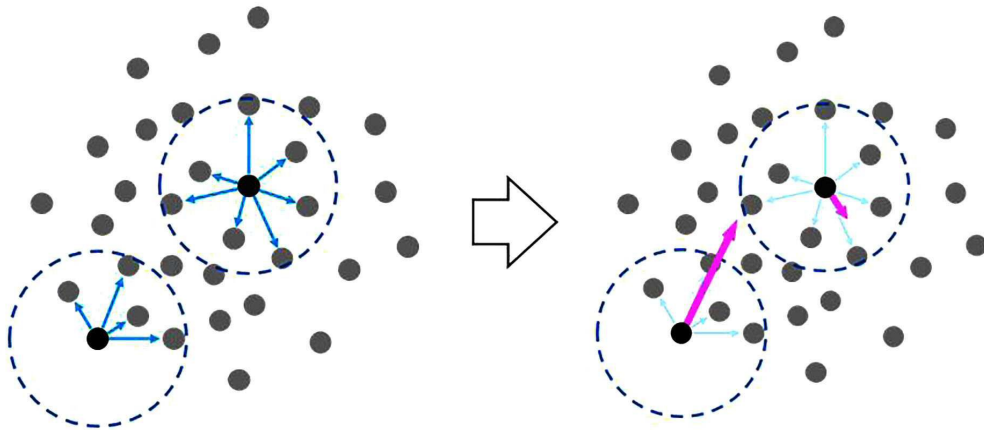
도면20



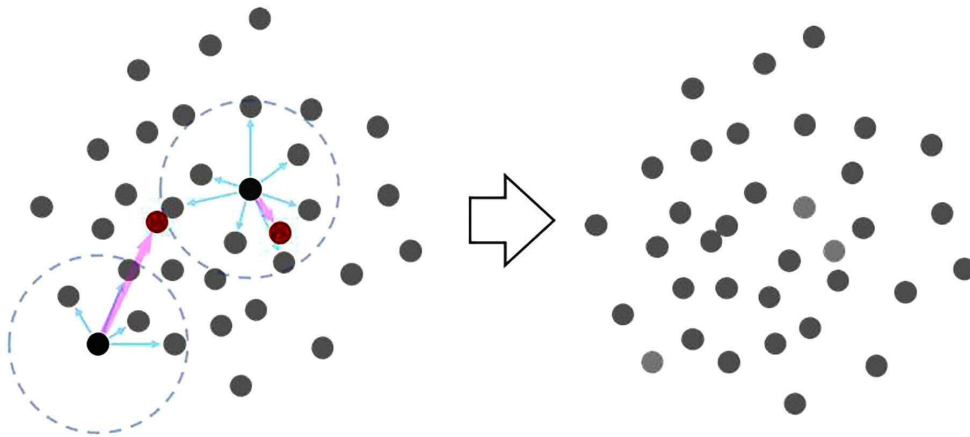
도면21



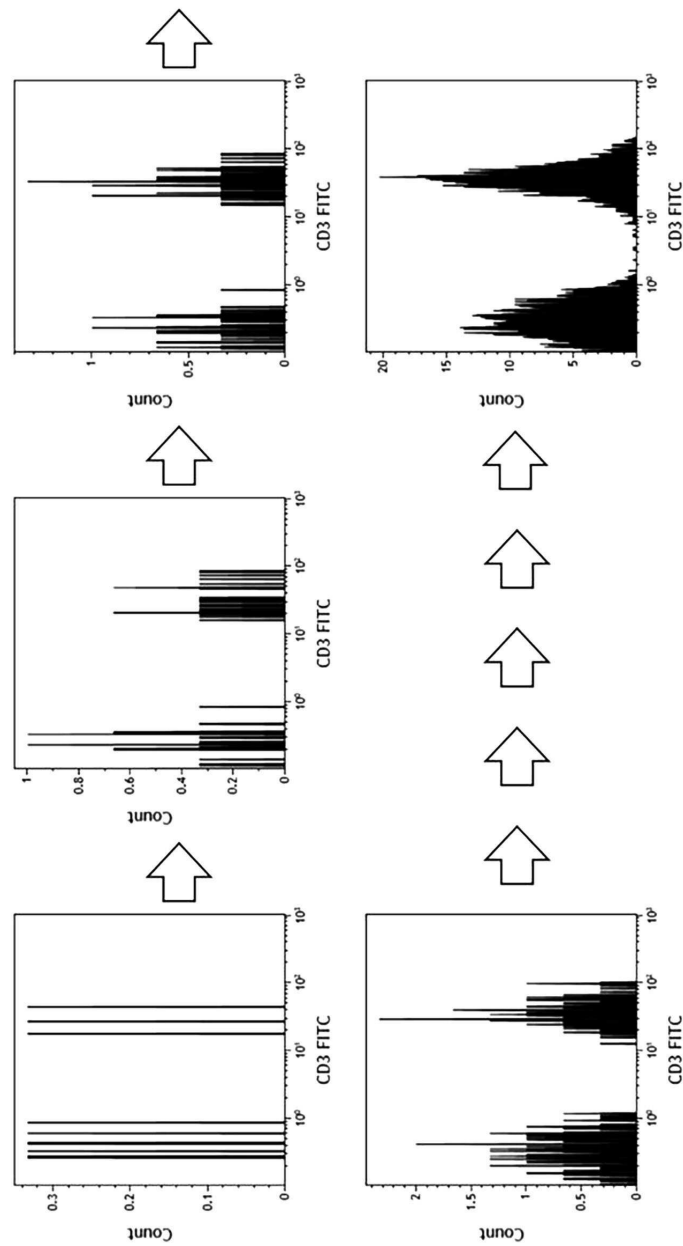
도면22



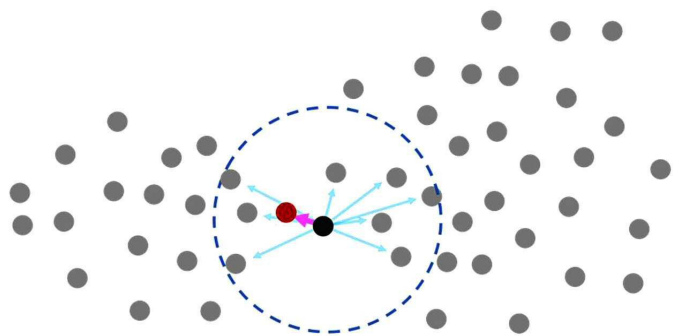
도면23



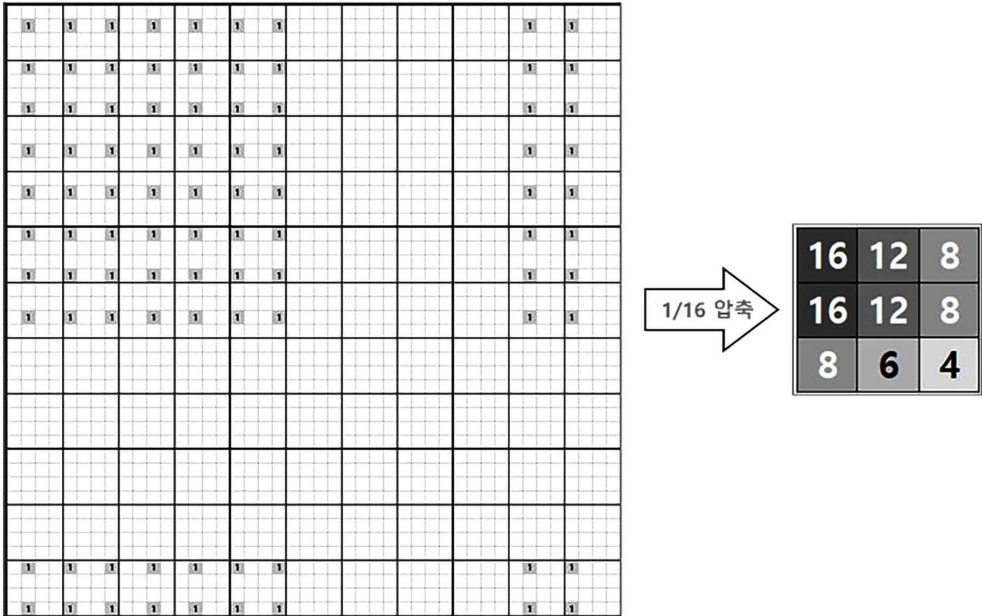
도면24



도면25



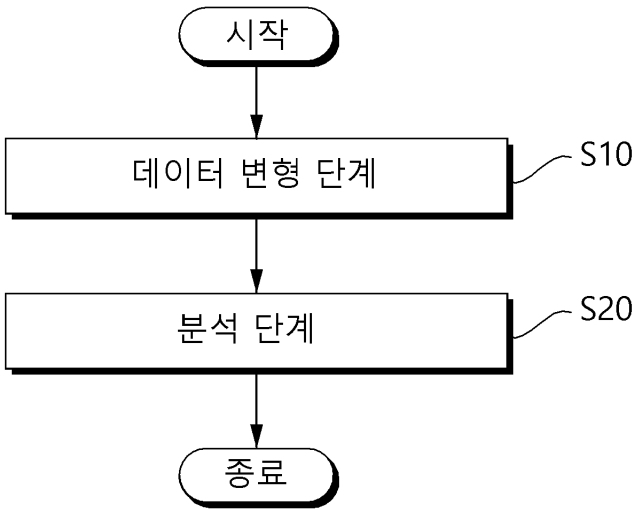
도면26



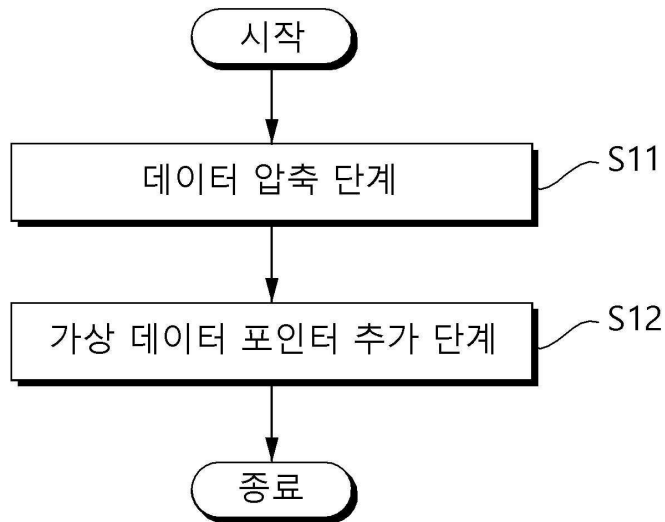
도면27

0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0	0
0	0	0	0	2	0	0	0	0	0	0	0
0	0	0	3	0	0	0	0	0	0	0	0
0	0	0	3	0	0	0	0	0	0	0	0
0	0	0	4	0	0	0	0	0	0	0	0
0	0	0	2	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0

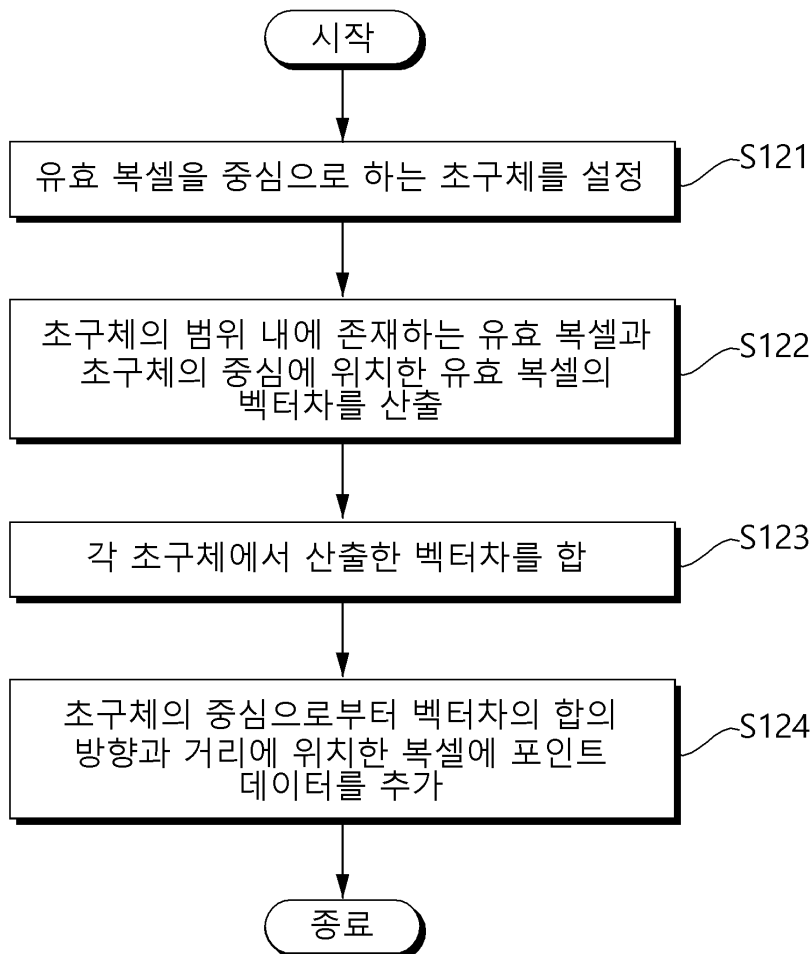
도면28



도면29



도면30



도면31

