



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년10월27일
(11) 등록번호 10-2317742
(24) 등록일자 2021년10월20일

- (51) 국제특허분류(Int. Cl.)
G06F 3/16 (2018.01) G06F 3/0484 (2013.01)
G06F 3/0488 (2013.01) G10L 15/30 (2013.01)
G10L 17/24 (2013.01)
- (52) CPC특허분류
G06F 3/167 (2013.01)
G06F 3/0484 (2013.01)
- (21) 출원번호 10-2020-0141831(분할)
- (22) 출원일자 2020년10월29일
심사청구일자 2020년11월27일
- (65) 공개번호 10-2020-0126357
- (43) 공개일자 2020년11월06일
- (62) 원출원 특허 10-2019-0157450
원출원일자 2019년11월29일
심사청구일자 2019년12월30일
- (56) 선행기술조사문헌
JP2013020411 A*
*는 심사관에 의하여 인용된 문헌

- (73) 특허권자
삼성전자 주식회사
경기도 수원시 영통구 삼성로 129 (매탄동)
- (72) 발명자
정지혜
경기도 수원시 영통구 삼성로 129 (매탄동)
신용욱
경기도 수원시 영통구 삼성로 129 (매탄동)
(뒷면에 계속)
- (74) 대리인
허성원, 이동욱, 서동현

전체 청구항 수 : 총 17 항

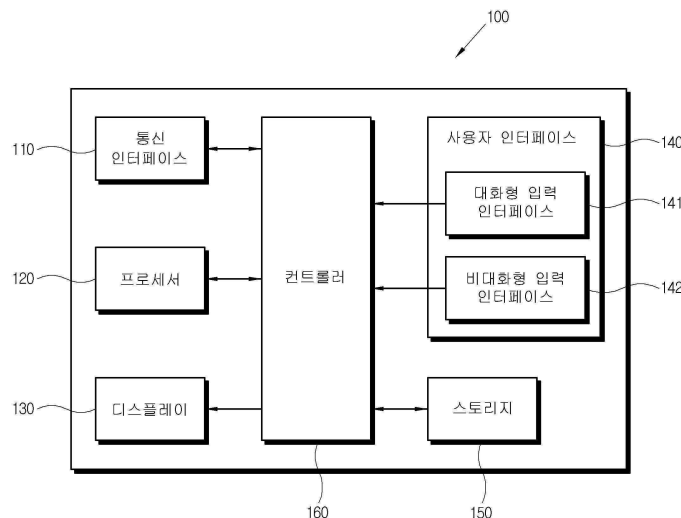
심사관 : 김건우

(54) 발명의 명칭 서버 및 그 제어방법과, 영상처리장치 및 그 제어방법

(57) 요약

본 발명의 실시예에 따른 서버는, 하나 이상의 영상처리장치와 통신하는 통신 인터페이스와; 데이터가 저장되는 스토리지와; 영상처리장치로부터 사용자의 발화에 의한 제1이벤트가 수신되면 제1이벤트에 대해 수행한 처리 결과를 영상처리장치에 제공하고, 제1이벤트의 처리에 따라서 제1이벤트의 이력을 스토리지에 저장하는 프로세서를 포함하며, 프로세서는, 영상처리장치로부터 비발화 방식의 사용자 입력에 의한 제2이벤트가 수신되면 제1이벤트 및 상기 제2이벤트 사이의 관련성을 판단하고, 제1이벤트 및 상기 제2이벤트가 관련성이 있다고 판단하면 스토리지에 저장된 제1이벤트의 이력에 기초하여 제2이벤트를 처리하는 것을 특징으로 한다.

대표도 - 도1



(52) CPC특허분류

G06F 3/0488 (2013.01)

G10L 15/30 (2013.01)

G10L 17/24 (2013.01)

(72) 발명자

신승민

경기도 수원시 영통구 삼성로 129 (매탄동)

이희란

경기도 수원시 영통구 삼성로 129 (매탄동)

최성욱

경기도 수원시 영통구 삼성로 129 (매탄동)

명세서

청구범위

청구항 1

전자장치에 있어서,

디스플레이부;

통신 인터페이스;

사용자 인터페이스; 및

상기 사용자 인터페이스를 통해 수신된 사용자 음성 입력에 기초하여, 제1 검색 요청을 서버로 전송하도록 상기 통신 인터페이스를 제어하고,

상기 통신 인터페이스를 통해 수신한, 상기 사용자 음성 입력에 대응하는, 제1 정보를 표시하도록 상기 디스플레이부를 제어하고,

상기 제1 정보를 표시하는 동안, 상기 사용자 인터페이스를 통해 수신된 사용자 입력에 기초하여, 상기 통신 인터페이스를 통해 수신한, 상기 사용자 입력에 대응하는, 제2 정보를 표시하도록 상기 디스플레이부를 제어하는 프로세서를 포함하고,

상기 제2 정보는, 상기 사용자 입력이 상기 사용자 음성 입력과 관련있는지에 따라, 상기 제1 정보와 관련된 정보 또는 상기 제1 정보와 독립된 정보인 것을 특징으로 하는 전자장치.

청구항 2

제1항에 있어서,

상기 제1 정보는 상기 사용자 음성 입력에 기초하여 검색된 적어도 하나의 검색 결과를 포함하는 전자장치.

청구항 3

제1항에 있어서,

상기 프로세서는, 상기 사용자 입력이 상기 제1 정보와 관련성이 있는지 여부를 식별하는 전자장치.

청구항 4

제1항에 있어서,

상기 사용자 입력은 입력 메뉴 또는 UI 이미지를 통해 상기 제2 정보를 검색하기 위한 터치 입력을 포함하는 전자장치.

청구항 5

제1항에 있어서,

저장부를 더 포함하고,

상기 프로세서는, 상기 사용자 음성 입력에 대응한 기능 수행의 이력 정보를 상기 저장부에 저장하는 전자장치.

청구항 6

제1항에 있어서,

상기 프로세서는, 상기 제2 정보를 검색하기 위한 제2 검색 요청이 상기 제1 검색 요청과 관련하여 처리할 수 있는 검색 요청인지에 대한 판단을 수행하는 전자장치.

청구항 7

제1항에 있어서,

상기 프로세서는, 상기 서버가 판단한 상기 제2 정보를 검색하기 위한 제2 검색 요청이 상기 제1 검색 요청과 관련하여 처리할 수 있는 검색 요청인지에 대한 정보를 수신하는 전자장치.

청구항 8

제1항에 있어서,

상기 프로세서는,

상기 사용자 음성 입력과 관련된 적어도 하나의 제1 검색 결과 및 키워드를 상기 서버로부터 수신하여 상기 디스플레이부에 표시하고,

상기 표시된 적어도 하나의 키워드 중 하나를 선택하는 상기 사용자 입력을 수신하는 전자장치.

청구항 9

제8항에 있어서,

상기 사용자 입력은 입력 메뉴 또는 UI 이미지를 통해 상기 제2 정보를 검색하기 위한 터치 입력을 포함하는 전자장치.

청구항 10

전자장치의 제어방법에 있어서,

사용자 인터페이스를 통해 수신된 사용자 음성 입력에 기초하여, 제1 검색 요청을 서버로 전송하는 단계;

통신 인터페이스를 통해 수신한, 상기 사용자 음성 입력에 대응하는, 제1 정보를 표시하는 단계; 및

상기 제1 정보를 표시하는 동안, 상기 사용자 인터페이스를 통해 수신된 사용자 입력에 기초하여, 상기 통신 인터페이스를 통해 수신한, 상기 사용자 입력에 대응하는 제2 정보를 표시하는 단계를 포함하고,

상기 제2 정보는 상기 사용자 입력이 상기 사용자 음성 입력과 관련있는지에 따라, 상기 제1 정보와 관련된 정보 또는 상기 제1 정보와 독립된 정보인 것을 특징으로 하는 전자장치의 제어방법.

청구항 11

제10항에 있어서,

상기 제1 정보는 상기 사용자 음성 입력에 기초하여 검색된 적어도 하나의 검색 결과를 포함하는 전자장치의 제어방법.

청구항 12

제10항에 있어서,

상기 사용자 입력이 상기 제1 정보와 관련성이 있는지 여부를 식별하는 단계를 포함하는 전자장치의 제어방법.

청구항 13

제10항에 있어서,

상기 사용자 입력은 입력 메뉴 또는 UI 이미지를 통해 상기 제2 정보를 검색하기 위한 터치 입력을 포함하는 전자장치의 제어방법.

청구항 14

제10항에 있어서,

상기 사용자 음성 입력에 대응한 기능 수행의 이력 정보를 저장부에 저장하는 단계를 포함하는 전자장치의 제어방법.

청구항 15

제10항에 있어서,

상기 제2 정보를 검색하기 위한 제2 검색 요청이 상기 제1 검색 요청과 관련하여 처리할 수 있는 검색 요청인지에 대한 판단을 수행하는 단계를 포함하는 전자장치의 제어방법.

청구항 16

제10항에 있어서,

상기 서버가 판단한 상기 제2 정보를 검색하기 위한 제2 검색 요청이 상기 제1 검색 요청과 관련하여 처리할 수 있는 검색 요청인지에 대한 정보를 수신하는 단계를 더 포함하는 전자장치의 제어방법.

청구항 17

제10항에 있어서,

상기 사용자 음성 입력과 관련된 적어도 하나의 제1 검색 결과 및 키워드를 상기 서버로부터 수신하여 표시하는 단계; 및

상기 표시된 적어도 하나의 키워드 중 하나를 선택하는 상기 사용자 입력을 수신하는 단계를 포함하는 전자장치의 제어방법.

발명의 설명

기술 분야

[0001] 본 발명은 영상신호를 영상으로 표시되게 처리하는 시스템 내의 서버 및 그 제어방법과, 영상처리장치 및 그 제어방법에 관한 것으로서, 상세하게는 사용자의 음성 명령을 인식함으로써 해당 음성 명령에 대응하는 기능 또는 동작을 실행할 수 있는 구조의 서버 및 그 제어방법과, 영상처리장치 및 그 제어방법에 관한 것이다.

배경 기술

[0002] 영상처리장치는 외부로부터 수신되는 영상신호/영상데이터를 다양한 영상처리 프로세스에 따라서 처리한다. 영상처리장치는 처리된 영상신호를 자체 구비한 디스플레이 패널 상에 영상으로 표시하거나, 또는 패널을 구비한 타 디스플레이장치에서 영상으로 표시되도록 이 처리된 영상신호를 해당 디스플레이장치에 출력할 수 있다. 즉, 영상처리장치는 영상신호를 처리 가능한 장치라면 영상을 표시 가능한 패널을 포함하는 경우 및 패널을 포함하지 않는 경우 모두 포함할 수 있는 바, 전자의 경우의 예시로는 TV가 있으며, 후자의 경우의 예시로는 셋탑박스(set-top box)가 있다.

[0003] 영상처리장치는 기술의 발전에 따라서 다양한 기능의 추가 및 확장이 계속적으로 반영되고 있는 바, 이러한 추세에 따라서 영상처리장치에 있어서 사용자의 의도를 반영한 커맨드를 영상처리장치에 입력하는 구성도 다양한 구조 또는 방법이 제안되고 있다. 예를 들면, 종래에는 사용자 리모트 컨트롤러(remote controller) 상의 키/버튼을 누르면 리모트 컨트롤러가 사용자가 원하는 동작이 실행되도록 하는 제어신호를 영상처리장치에 전송하는 구성이었으나, 근래에는 영상처리장치가 사용자에게 의한 모션 또는 발화 등을 감지하고, 감지된 내용을 분석하여 대응 동작을 실행시키는 등, 사용자의 의도를 반영하여 영상처리장치를 제어하는 다양한 구성이 제안되고 있다.

발명의 내용

[0004] 본 발명의 실시예에 따른 서버는, 하나 이상의 영상처리장치와 통신하는 통신 인터페이스와; 데이터가 저장되는 스토리지와; 상기 영상처리장치로부터 사용자의 발화에 의한 제1이벤트가 수신되면 상기 제1이벤트에 대해 수행한 처리 결과를 상기 영상처리장치에 제공하고, 상기 제1이벤트의 처리에 따라서 상기 제1이벤트의 이력을 상기 스토리지에 저장하는 프로세서를 포함하며, 상기 프로세서는, 상기 영상처리장치로부터 비발화 방식의 사용자 입력에 의한 제2이벤트가 수신되면 상기 제1이벤트 및 상기 제2이벤트 사이의 관련성을 판단하고, 상기 제1이벤트 및 상기 제2이벤트가 관련성이 있다고 판단하면 상기 스토리지에 저장된 상기 제1이벤트의 이력에 기초하여 상기 제2이벤트를 처리하는 것을 특징으로 한다.

[0005] 여기서, 상기 프로세서는, 상기 제2이벤트가 수신되면, 상기 스토리지에 저장된 상기 제1이벤트의 이력을 참조

하여, 상기 제1이벤트와 독립적으로 상기 제2이벤트를 처리 가능한지 아니면 상기 제1이벤트의 처리 결과에 기초하여 상기 제2이벤트를 처리하는 것인지의 여부를 판단할 수 있다.

- [0006] 또한, 상기 프로세서는, 상기 제1이벤트가 수신되면 상기 제1이벤트에 기초하여 기 설정된 데이터베이스에서 검색을 수행하고, 상기 제1이벤트로부터 복수의 상기 검색 결과가 도출된 상태에서 상기 제2이벤트가 수신되면, 상기 제1이벤트에 대한 상기 복수의 검색 결과 중에서 상기 제2이벤트에 기초하여 선택된 하나의 최종 결과를 상기 영상처리장치에 전송할 수 있다.
- [0007] 여기서, 상기 프로세서는, 상기 제1이벤트에 대한 상기 검색 결과가 복수 개이면, 상기 복수의 검색 결과와, 상기 복수의 검색 결과 중 어느 하나의 선택을 요청하는 요청 커맨드를 상기 영상처리장치에 전송하며, 상기 영상처리장치로부터 상기 요청 커맨드에 대응하는 상기 제2이벤트를 수신할 수 있다.
- [0008] 또한, 상기 제2이벤트는 상기 서버가 인식 가능하도록 상기 영상처리장치에 의해 발화에 따른 음성 명령의 형태로 변환되어 상기 서버에 수신될 수 있다.
- [0009] 또한, 본 발명의 실시예에 따른 서버의 제어방법은, 영상처리장치로부터 사용자의 발화에 의한 제1이벤트를 수신하는 단계와; 상기 제1이벤트에 대해 수행한 처리 결과를 상기 영상처리장치에 제공하고, 상기 제1이벤트의 처리에 따라서 상기 제1이벤트의 이력을 저장하는 단계와; 상기 영상처리장치로부터 비발화 방식의 사용자 입력에 의한 제2이벤트를 수신하는 단계와; 상기 제1이벤트 및 상기 제2이벤트 사이의 관련성을 판단하고, 상기 제1이벤트 및 상기 제2이벤트가 관련성이 있다고 판단하면 기 저장된 상기 제1이벤트의 이력에 기초하여 상기 제2이벤트를 처리하는 단계를 포함하는 것을 특징으로 한다.
- [0010] 여기서, 상기 제1이벤트 및 상기 제2이벤트 사이의 관련성을 판단하는 단계는, 상기 기 저장된 상기 제1이벤트의 이력을 참조하여, 상기 제1이벤트와 독립적으로 상기 제2이벤트를 처리 가능한지 아니면 상기 제1이벤트의 처리 결과에 기초하여 상기 제2이벤트를 처리하는 것인지의 여부를 판단하는 단계를 포함할 수 있다.
- [0011] 또한, 상기 제1이벤트에 대해 수행한 처리 결과를 상기 영상처리장치에 제공하는 단계는, 상기 제1이벤트에 기초하여 기 설정된 데이터베이스에서 검색을 수행하는 단계를 포함하며, 상기 기 저장된 상기 제1이벤트의 이력에 기초하여 상기 제2이벤트를 처리하는 단계는, 상기 제1이벤트로부터 복수의 상기 검색 결과가 도출된 상태에서 상기 제2이벤트가 수신되면, 상기 제1이벤트에 대한 상기 복수의 검색 결과 중에서 상기 제2이벤트에 기초하여 선택된 하나의 최종 결과를 상기 영상처리장치에 전송하는 단계를 포함할 수 있다.
- [0012] 여기서, 상기 제2이벤트를 수신하는 단계는, 상기 제1이벤트에 대한 상기 검색 결과가 복수 개이면, 상기 복수의 검색 결과와, 상기 복수의 검색 결과 중 어느 하나의 선택을 요청하는 요청 커맨드를 상기 영상처리장치에 전송하는 단계와; 상기 영상처리장치로부터 상기 요청 커맨드에 대응하는 상기 제2이벤트를 수신하는 단계를 포함할 수 있다.
- [0013] 또한, 상기 제2이벤트는 상기 서버가 인식 가능하도록 상기 영상처리장치에 의해 발화에 따른 음성 명령의 형태로 변환되어 상기 서버에 수신될 수 있다.
- [0014] 또한, 본 발명의 실시예에 따른 영상처리장치는, 데이터가 저장되는 스토리지와; 사용자의 발화가 입력되는 발화 입력 인터페이스와, 발화가 아닌 방식의 사용자 입력이 수행되는 비발화 입력 인터페이스를 포함하는 사용자 인터페이스와; 상기 발화 입력 인터페이스에 의해 발생하는 제1이벤트에 대응하는 동작이 실행되도록 처리하는 프로세서와; 상기 제1이벤트의 발생 및 처리에 대한 이력을 상기 스토리지에 저장하며, 상기 제1이벤트의 발생 이후에 상기 비발화 입력 인터페이스에 의해 제2이벤트가 발생하면 상기 제1이벤트 및 상기 제2이벤트 사이의 관련성을 판단하고, 상기 제1이벤트 및 상기 제2이벤트가 관련성이 있다고 판단하면 상기 스토리지에 저장된 상기 제1이벤트의 이력에 기초하여 상기 제2이벤트가 처리되게 제어하는 컨트롤러를 포함하는 것을 특징으로 한다.
- [0015] 여기서, 상기 컨트롤러는, 상기 제2이벤트가 발생하면, 상기 스토리지에 저장된 상기 제1이벤트의 이력을 참조하여, 상기 제1이벤트와 독립적으로 상기 제2이벤트를 처리 가능한지 아니면 상기 제1이벤트의 처리 결과에 기초하여 상기 제2이벤트를 처리하는 것인지의 여부를 판단할 수 있다.
- [0016] 또한, 상기 컨트롤러는, 상기 제1이벤트에 대응하는 동작에 대해 복수의 결과가 검색되면 상기 복수의 검색 결과를 선택 가능하게 디스플레이에 표시하고, 상기 복수의 검색 결과가 표시된 상태에서 상기 제2이벤트가 발생하면 상기 제2이벤트가 상기 복수의 검색 결과 중 어느 하나의 선택을 지시하는지 여부를 판단할 수 있다.
- [0017] 여기서, 발화에 대응하는 음성 명령에 대한 검색 처리를 수행하는 서버와 통신하는 통신 인터페이스를 더 포함

하며, 상기 컨트롤러는, 상기 제1이벤트를 상기 서버에 전송하며 상기 제1이벤트에 대한 검색 결과를 상기 서버로부터 수신하여 디스플레이에 표시할 수 있다.

[0018] 또한, 상기 컨트롤러는, 상기 서버가 상기 제2이벤트를 인식 가능하도록 상기 제2이벤트를 발화에 따른 음성 명령의 형태로 변환하여 상기 서버에 전송할 수 있다.

[0019] 또한, 본 발명의 실시예에 따른 영상처리장치의 제어방법은, 사용자의 발화가 입력되는 발화 입력 인터페이스에 의해 제1이벤트가 발생하는 단계와; 상기 제1이벤트에 대응하는 동작을 실행하는 단계를 포함하며, 상기 제1이벤트에 대응하는 동작을 실행하는 단계는, 상기 제1이벤트의 발생 및 처리에 대한 이력을 상기 스토리지에 저장하는 단계와; 발화가 아닌 방식의 사용자 입력이 수행되는 비발화 입력 인터페이스에 의해 제2이벤트가 발생하는 단계와; 상기 제1이벤트 및 상기 제2이벤트 사이의 관련성을 판단하고, 상기 제1이벤트 및 상기 제2이벤트가 관련성이 있다고 판단하면 상기 저장된 제1이벤트의 이력에 기초하여 상기 제2이벤트가 처리되게 제어하는 단계를 포함하는 것을 특징으로 한다.

[0020] 여기서, 상기 상기 제1이벤트 및 상기 제2이벤트 사이의 관련성을 판단하는 단계는, 상기 기 저장된 제1이벤트의 이력을 참조하여, 상기 제1이벤트와 독립적으로 상기 제2이벤트를 처리 가능한지 아니면 상기 제1이벤트의 처리 결과에 기초하여 상기 제2이벤트를 처리하는 것인지의 여부를 판단하는 단계를 포함할 수 있다.

[0021] 또한, 상기 상기 제1이벤트에 대응하는 동작을 실행하는 단계는, 상기 제1이벤트에 대응하는 동작에 대해 복수의 결과가 검색되면 상기 복수의 검색 결과를 선택 가능하게 표시하는 단계를 더 포함하며, 상기 제1이벤트 및 상기 제2이벤트 사이의 관련성을 판단하는 단계는, 상기 제2이벤트가 상기 복수의 검색 결과 중 어느 하나의 선택을 지시하는지 여부를 판단하는 단계를 포함할 수 있다.

[0022] 여기서, 상기 영상처리장치는 발화에 대응하는 음성 명령에 대한 검색 처리를 수행하는 서버와 통신하며, 상기 제1이벤트를 상기 서버에 전송함으로써 상기 제1이벤트에 대한 검색 결과를 상기 서버로부터 수신하여 표시할 수 있다.

[0023] 또한, 상기 영상처리장치는, 상기 서버가 상기 제2이벤트를 인식 가능하도록 상기 제2이벤트를 발화에 따른 음성 명령의 형태로 변환하여 상기 서버에 전송할 수 있다.

[0024] 또한, 본 발명의 실시예에 따른 영상처리장치는, 사용자의 발화가 입력되는 발화 입력 인터페이스와, 발화가 아닌 방식의 사용자 입력이 수행되는 비발화 입력 인터페이스를 포함하는 사용자 인터페이스와; 발화에 의한 음성 명령에 대응하는 동작을 검색하는 서버와 통신하는 통신 인터페이스와; 상기 발화 입력 인터페이스에 의해 발생하는 제1이벤트에 대한 검색 결과를 상기 서버로부터 제공받고, 상기 검색 결과가 실행되도록 처리하는 프로세서와; 상기 서버에 의해 상기 제1이벤트에 대하여 복수의 상기 검색 결과가 도출된 상태에서 상기 비발화 입력 인터페이스에 의해 상기 복수의 검색 결과 중 어느 하나의 선택을 지시하는 제2이벤트가 발생하면, 상기 서버가 상기 제2이벤트에 기초하여 상기 복수의 검색 결과 중 어느 하나를 선택하여 최종 검색 결과를 제공할 수 있도록, 상기 제2이벤트를 상기 서버가 인식 가능한 음성 명령의 형태로 변환하여 상기 서버에 전송하는 컨트롤러를 포함하는 것을 특징으로 한다.

도면의 간단한 설명

- [0025] 도 1은 본 발명의 제1실시예에 따른 시스템이 포함하는 영상처리장치의 구성 블록도,
- 도 2는 본 발명의 제1실시예에 따른 시스템이 포함하는 서버의 구성 블록도,
- 도 3 및 도 4는 도 2의 대화형 서버가 이벤트에 대응하는 검색 결과를 디스플레이장치에 제공하는 방법을 나타내는 플로우차트,
- 도 5는 도 2의 대화형 서버의 서버측 프로세서를 기능별로 세분화시킨 구성 블록도,
- 도 6 및 도 7은 도 2의 디스플레이장치 및 대화형 서버와, 대화형 서버 내의 각 구성요소들 사이에서 수행되는 데이터의 전송 관계를 나타내는 예시도,
- 도 8 및 도 9는 본 발명의 제2실시예에 따른 디스플레이장치의 제어방법을 나타내는 플로우차트,
- 도 10은 본 발명의 제3실시예에 따른 디스플레이장치의 프로세서의 구성 블록도이다.

발명을 실시하기 위한 구체적인 내용

- [0026] 이하에서는 첨부도면을 참조하여 본 발명에 대해 상세히 설명한다. 이하 실시예에서는 본 발명의 사상과 직접적인 관련이 있는 구성들에 대해서만 설명하며, 그 외의 구성에 대해서는 설명을 생략한다. 그러나, 본 발명의 사상이 적용된 장치 또는 시스템을 구현함에 있어서, 이와 같이 설명이 생략된 구성이 불필요함을 의미하는 것이 아님을 밝힌다.
- [0027] 도 1은 본 발명의 제1실시예에 따른 시스템이 포함하는 영상처리장치(100)의 구성 블록도이다.
- [0028] 이하 실시예는 영상처리장치(100)가 자체적으로 영상을 표시할 수 있는 구조의 디스플레이장치, 예를 들면 TV로 구현된 경우에 대해 설명한다. 그러나, 본 발명의 사상은 영상처리장치(100)가 자체적으로 영상을 표시하지 않고 타 디스플레이장치에 영상데이터 또는 제어정보를 출력 가능한 구조의 장치인 경우에도 적용이 가능한 바, 본 발명의 구현 형태가 이하 설명하는 실시예에 한정되지 않는다. 즉, 본 발명의 사상은 다양하게 변경되어 적용됨으로써, 다양한 방식의 장치에서 구현될 수 있다.
- [0029] 도 1에 도시된 바와 같이, 본 실시예에 따른 영상처리장치(100) 또는 디스플레이장치(100)는 외부의 영상공급원(미도시)로부터 영상신호를 수신한다. 디스플레이장치(100)가 수신 가능한 영상신호는 그 종류 또는 특성이 한정되지 않는 바, 예를 들면 디스플레이장치(100)는 방송국의 송출장비(미도시)로부터 송출되는 방송신호를 수신하고, 해당 방송신호를 튜닝하여 방송영상을 표시할 수 있다.
- [0030] 디스플레이장치(100)는 외부와 데이터/신호를 송수신하도록 통신하는 통신 인터페이스(interface)(110)와, 통신 인터페이스(110)에 수신되는 데이터를 기 설정된 프로세스(process)에 따라서 처리하는 프로세서(processor)(120)와, 프로세서(120)에서 처리되는 데이터가 영상데이터인 경우에 해당 영상데이터를 영상으로 표시하는 디스플레이(display)(130)와, 사용자에게 의한 입력 동작이 수행되는 사용자 인터페이스(140)와, 데이터/정보가 저장되는 스토리지(storage)(150)와, 디스플레이장치(100)의 제반 동작을 제어하는 컨트롤러(controller)(160)를 포함한다.
- [0031] 통신 인터페이스(110)는 디스플레이장치(100)가 서버 등과 같은 외부장치(미도시)와 양방향 통신이 가능하도록 데이터의 송수신을 수행한다. 통신 인터페이스(110)는 기 설정된 통신 프로토콜(protocol)에 따라서, 유선/무선을 통한 광역/근거리 네트워크 또는 로컬 접속 방식으로 외부장치(미도시)에 접속한다.
- [0032] 통신 인터페이스(110)는 각 장치 별 접속 포트(port) 또는 접속 모듈(module)의 집합체에 의해 구현될 수 있는 바, 접속을 위한 프로토콜 또는 접속 대상이 되는 외부장치가 하나의 종류 또는 형식으로 한정되지 않는다. 통신 인터페이스(110)는 디스플레이장치(100)에 내장된 형태일 수 있으며, 전체 또는 일부의 구성이 디스플레이장치(100)에 애드-온(add-on) 또는 dongle 형태로 추가 설치되는 형식도 가능하다.
- [0033] 통신 인터페이스(110)는 접속된 각 장치 별로 지정된 프로토콜에 따라서 신호의 송수신이 이루어지는 바, 접속된 각 장치에 대해 개별적인 접속 프로토콜 기반으로 신호를 송수신할 수 있다. 영상데이터의 경우를 예로 들면, 통신 인터페이스(110)는 RF(radio frequency)신호, 콤포지트(composite)/컴포넌트(component) 비디오, 슈퍼 비디오(super video), SCART, HDMI(high definition multimedia interface), 디스플레이포트(DisplayPort), UDI(unified display interface), 또는 와이어리스(wireless) HD 등 다양한 규격을 기반으로 신호의 송수신이 가능하다.
- [0034] 프로세서(120)는 통신 인터페이스(110)에 수신되는 데이터/신호에 대해 다양한 프로세스를 수행한다. 통신 인터페이스(110)에 영상데이터가 수신되면, 프로세서(120)는 영상데이터에 대해 영상처리 프로세스를 수행하고, 이러한 프로세스가 수행된 영상데이터를 디스플레이(130)에 출력함으로써 디스플레이(130)에 해당 영상데이터에 기초하는 영상이 표시되게 한다. 통신 인터페이스(110)에 수신되는 신호가 방송신호인 경우, 프로세서(120)는 특정 채널로 튜닝된 방송신호로부터 영상, 음성 및 부가데이터를 추출하고, 영상을 기 설정된 해상도로 조정하여 디스플레이(130)에 표시되게 한다.
- [0035] 프로세서(120)가 수행하는 영상처리 프로세스의 종류는 한정되지 않으며, 예를 들면 영상데이터의 영상 포맷에 대응하는 디코딩(decoding), 인터레이스(interlace) 방식의 영상데이터를 프로그레시브(progressive) 방식으로 변환하는 디인터레이싱(de-interlacing), 영상데이터를 기 설정된 해상도로 조정하는 스케일링(scaling), 영상 화질 개선을 위한 노이즈 감소(noise reduction), 디테일 강화(detail enhancement), 프레임 리프레시 레이트(frame refresh rate) 변환 등을 포함할 수 있다.
- [0036] 프로세서(120)는 데이터의 종류, 특성에 따라서 다양한 프로세스를 수행할 수 있는 바, 프로세서(120)가 수행 가능한 프로세스를 영상처리 프로세스로 한정할 수 없다. 또한, 프로세서(120)가 처리 가능한 데이터가 통신 인

터페이스(110)에 수신되는 것만으로 한정할 수 없다. 예를 들면, 프로세서(120)는 사용자 인터페이스(140)를 통해 사용자의 발화가 입력되면, 기 설정된 음성처리 프로세스에 따라서 해당 발화를 처리할 수 있다.

- [0037] 프로세서(120)는 이러한 여러 기능을 통합시킨 SOC(system-on-chip)가, 또는 이러한 각 프로세스를 독자적으로 수행할 수 있는 개별적인 칩셋(chip-set)들이 인쇄회로기판 상에 장착된 영상처리부드(미도시)로 구현되며, 디스플레이장치(100)에 내장된다.
- [0038] 디스플레이(130)는 프로세서(120)에 의해 처리된 영상신호/영상데이터를 영상으로 표시한다. 디스플레이(130)의 구현 방식은 한정되지 않는 바, 액정(liquid crystal), 플라즈마(plasma), 발광 다이오드(light-emitting diode), 유기발광 다이오드(organic light-emitting diode), 면전도 전자총(surface-conduction electron-emitter), 탄소 나노 튜브(carbon nano-tube), 나노 크리스탈(nano-crystal) 등의 다양한 디스플레이 방식으로 구현될 수 있다.
- [0039] 디스플레이(130)는 그 구현 방식에 따라서 부가적인 구성을 추가적으로 포함할 수 있다. 예를 들면, 액정 방식인 경우에, 디스플레이(130)는 액정 디스플레이 패널(미도시)과, 이에 광을 공급하는 백라이트유닛(미도시)과, 패널(미도시)을 구동시키는 패널구동기판(미도시) 등을 포함한다.
- [0040] 사용자 인터페이스(140)는 사용자의 조작 또는 입력에 따라서 기 설정된 다양한 제어 커맨드(command) 또는 정보를 컨트롤러(160)에 전달한다. 사용자 인터페이스(140)는 사용자의 의도에 따라서 사용자에게 의해 발생하는 다양한 이벤트(event)를 정보화하여 컨트롤러(160)에 전달한다. 여기서, 사용자에게 의해 발생하는 이벤트는 여러 가지 형태가 가능한 바, 예를 들면 사용자의 조작, 발화, 제스처(gesture) 등이 가능하다.
- [0041] 사용자 인터페이스(140)는 사용자가 정보를 입력하는 방식에 대응하여, 해당 방식으로 입력되는 정보를 감지할 수 있게 마련된다. 이에 따라서, 사용자 인터페이스(140)는 대화형 입력 인터페이스(141) 및 비대화형 입력 인터페이스(142)로 구별할 수 있다.
- [0042] 대화형 입력 인터페이스(141)는 사용자의 발화가 입력되도록 마련된다. 즉, 대화형 입력 인터페이스(141)는 마이크로폰으로 구현되며, 디스플레이장치(100)의 외부 환경에서 발생하는 다양한 소리를 감지한다. 대화형 입력 인터페이스(141)는 일반적으로 사용자에게 의한 발화를 감지할 수 있지만, 사용자 이외의 다양한 환경적 요인에 의해 발생하는 소리도 감지할 수 있다.
- [0043] 비대화형 입력 인터페이스(142)는 발화 방식 이외의 사용자의 입력이 수행되게 마련된다. 이 경우에 비대화형 입력 인터페이스(142)의 구현 형태는 다양하며, 예를 들면, 디스플레이장치(100)와 분리 이격된 리모트 컨트롤러(remote controller)나, 디스플레이장치(100) 외측에 설치된 메뉴 키(menu-key) 또는 입력 패널(panel)이나, 사용자의 제스처를 감지하는 모션센서(motion sensor) 또는 카메라(camera) 등이 가능하다.
- [0044] 또는, 비대화형 입력 인터페이스(142)는 디스플레이(130)에 설치된 터치스크린(touch-screen)으로 구현될 수도 있다. 이 경우에, 사용자는 디스플레이(130)에 표시된 입력메뉴 또는 UI 영상을 터치함으로써, 기 설정된 커맨드 또는 정보가 컨트롤러(160)에 전달되게 할 수 있다.
- [0045] 스토리지(150)는 컨트롤러(160)의 제어에 따라서 다양한 데이터가 저장된다. 스토리지(150)는 시스템 전원의 제공 유무와 무관하게 데이터를 보존할 수 있도록, 플래시메모리(flash-memory), 하드디스크 드라이브(hard-disc drive)와 같은 비휘발성 메모리로 구현된다. 스토리지(150)는 컨트롤러(160)에 대해 액세스(access)됨으로써, 기 저장된 데이터의 독취, 기록, 수정, 삭제, 갱신 등이 수행된다.
- [0046] 컨트롤러(160)는 CPU로 구현되며, 소정 이벤트의 발생에 따라서 프로세서(120)를 비롯한 디스플레이장치(100)의 제반 구성의 동작을 제어한다. 사용자 발화 이벤트의 경우를 예를 들면, 컨트롤러(160)는 대화형 입력 인터페이스(141)를 통해 사용자의 발화가 입력되면 이 입력된 발화를 처리하도록 프로세서(120)를 제어한다. 만일 사용자가 채널 번호를 발화하였다면, 컨트롤러(160)는 발화한 채널번호로 방송영상이 전환 표시되도록 제어한다.
- [0047] 여기서, 입력된 발화를 처리하는 과정은 디스플레이장치(100) 단독으로 수행할 수도 있지만, 시스템의 부하나 처리능력 등 여러 요인을 고려하여 외부장치와 분담하여 수행할 수도 있다. 이러한 구성에 관하여 이하 설명한다.
- [0048] 도 2는 본 실시예의 시스템이 포함하는 서버(200, 300)의 구성 블록도이다.
- [0049] 도 2에 도시된 바와 같이, 디스플레이장치(100)는 네트워크를 통해 하나 이상의 서버(200, 300)와 통신 가능하게 접속한다. 여기서, 디스플레이장치(100)는 사용자의 발화를 음성 명령으로 변환하는 STT(speech-to-text) 서

버(200)와, 음성 명령을 분석함으로써 해당 음성 명령에 대응하는 동작/액션(action)/결과를 판단하는 대화형 서버(300)에 각기 접속한다. 여기서 각 서버(200, 300)의 명칭은 각 기능을 고려하여 단지 상호 구분을 위해 적용된 것인 바, 해당 명칭이 각 서버(200, 300)의 구성적 특징을 한정하지는 않는다.

- [0050] STT 서버(200)는 음성신호의 파형을 분석함으로써 음성신호의 내용을 텍스트로 생성한다. 이에 따라서, STT 서버(200)는 디스플레이장치(100)로부터 음성신호가 수신되면, 이를 텍스트의 음성 명령으로 변환한다. 여기서, 발화는 사용자에게 의해 이루어진 동작을 지칭하며, 음성 명령은 발화의 내용이 텍스트 형식으로 데이터화된 것을 지칭한다.
- [0051] 대화형 서버(300)는 음성 명령에 대응하는 다양한 디스플레이장치(100)의 동작을 검색한다. 대화형 서버(300)는 디스플레이장치(100)로부터 수신된 음성 명령을 분석하고, 분석 결과에 따라서 해당 음성 명령에 대응하는 동작을 수행하기 위한 제어신호나 또는 해당 음성 명령에 기초하여 검색된 결과 정보를 디스플레이장치(100)에 전송한다.
- [0052] 대화형 서버(300)는 디스플레이장치(100)와 통신하기 위한 서버측 통신 인터페이스(310)와, 서버측 통신 인터페이스(310)를 통해 수신되는 음성 명령에 대한 처리를 수행하는 서버측 프로세서(320)를 포함한다.
- [0053] 이러한 디스플레이장치(100), STT 서버(200) 및 대화형 서버(300)에 의한 시스템 하에서, 디스플레이장치(100)의 사용자가 발화하는 경우에 수행되는 동작에 관해 이하 설명한다.
- [0054] 디스플레이장치(100)는 사용자의 발화가 입력되면, 해당 발화에 의한 음성신호를 STT 서버(200)에 전송한다. STT 서버(200)는 디스플레이장치(100)로부터 수신되는 음성신호를 데이터화된 음성 명령으로 변환하고, 이 음성 명령을 디스플레이장치(100)에 전송한다. 디스플레이장치(100)는 STT 서버(200)로부터 수신된 음성 명령을 대화형 서버(300)로 전송한다.
- [0055] 대화형 서버(300)는 음성 명령의 내용을 분석하고, 음성 명령에 대응하는 결과를 도출하여 디스플레이장치(100)에 반환한다. 예를 들면, 음성 명령의 내용이 "이번 주의 'program ABC'의 출연자는 누구인가?"이라고 할 때, 대화형 서버(300)는 내부 또는 외부의 데이터베이스에서 "이번 주", "program ABC", "출연자"의 키워드를 가지고 검색을 수행하여 그 결과를 취득한다.
- [0056] 여기서, 데이터베이스는 키워드의 카테고리 별로 구조화(structured)되어 있으므로, 특정 키워드를 가지고 원하는 결과를 검색할 수 있다. 구조화된 데이터베이스는 데이터베이스의 설계 및 구축 방식에 따라서 다양한 방법으로 구현이 가능한 사항인 바, 이에 관한 자세한 설명은 생략한다.
- [0057] 그런데, 대화형 서버(300)가 음성 명령으로부터 도출한 하나 이상의 키워드로 검색을 수행했을 때, 복수의 검색 결과가 나올 수 있다. 앞선 음성 명령의 예를 들면, 이번 주에 방송하는 방송 프로그램인 "program ABC"가 각기 다른 시간 또는 각기 다른 방송채널에서 방송하는 경우를 생각할 수 있다.
- [0058] 이 경우에 대화형 서버(300)는 음성 명령에 대한 복수의 검색 결과 중에서, 사용자의 의도에 맞는 어느 하나의 검색 결과를 선택할 필요가 있다. 대화형 서버(300)가 하나의 검색 결과를 선택하는 대응 방법으로는, 예를 들면 다음과 같은 두 가지 방법이 가능하다.
- [0059] 첫 번째 방법은, 대화형 서버(300)의 클라이언트, 즉 디스플레이장치(100)가 자체적으로 수립한 특정 규칙에 의해 어느 하나의 검색 결과가 선택된다. 예를 들면, 음성 명령이 특정 프로그램의 방송 시간을 문의하는 내용이고, 이러한 조건을 만족하는 프로그램이 복수 개라면, 가장 최근에 방송될 프로그램이 선택될 수 있다. 두 번째 방법은, 대화형 서버(300)는 디스플레이장치(100)의 사용자가 어느 하나의 검색 결과를 선택하는 발화를 하면, 해당 발화를 이전 발화 이력과 연결시켜 사용자의 원래 의도에 맞는 검색 결과를 선택할 수 있다.
- [0060] 그런데, 이와 같은 시스템에서 사용자에게는 다양한 입력 방식의 환경이 제공되므로, 디스플레이장치(100)가 음성인식 입력 환경을 지원한다고 하더라도 사용자는 다양한 입력 방식으로 각 상황에서 가장 편리하다고 생각되는 방식을 선택하여 입력을 수행할 것이다. 이러한 다중 입력 모드를 지원하는 시스템의 환경에서, 만일 사용자의 다중 입력들 사이에 이력 관리를 개별적으로 하게 된다면, 사용자의 원래 의도를 반영하는 결과를 선택 또는 도출하는 것은 제약이 있을 것이다. 만일, 사용자가 최초에 발화 입력을 수행하고 이후에 다른 방식으로 입력 방법을 바꾸어 사용했다고 해서 사용자가 원래 의도한 결과가 제공되지 않는다면, 사용자는 사용 환경에 대해 불편을 느낄 것이다.
- [0061] 따라서, 대화형 서버(300) 및 디스플레이장치(100)에 의한 대화형 시스템 환경에 있어서, 사용자로부터의 발화에 의해 최초 검색된 결과가 복수 개이고 이후 사용자가 비발화 방식의 입력을 통해 어느 하나의 검색 결과를

선택하면, 대화형 시스템은 사용자가 최초 발화의 의도에 맞는 검색 결과를 제공할 필요가 있다.

- [0062] 여기서, 종래 기술에 따른 시스템에서는 이러한 다중 입력 모드 각각에 대한 이력 관리를 통합적으로 지원하지 않는다. 즉, 종래 기술에 따른 대화형 서버(300)에서는 디스플레이장치(100)의 사용자의 발화에 의한 입력과, 사용자의 제스처에 의한 입력과, 사용자의 리모트 컨트롤러(미도시)의 조작에 의한 입력을 개별적으로 관리하며, 이들 입력들 각각을 상호 연계시켜서 관리하지 않는다.
- [0063] 이러한 종래 기술에 따른 시스템에서, 앞서 두 번째 방법에 따라서 디스플레이장치(100)가 대화형 서버(300)로부터 수신한 복수의 검색 결과를 표시한 상태를 고려한다. 이 때, 사용자가 만일 발화가 아닌 비대화형 입력 인터페이스(142)를 통한 비발화 방식으로 입력을 수행할 수도 있는데, 종래 기술에 따른 대화형 서버(300)는 발화 입력 방식 및 비발화 입력 방식 간에 입력의 이력을 상호 연계시켜 참조하지 않으므로 사용자의 의도에 맞지 않는 결과가 나오게 된다.
- [0064] 예를 들어 사용자가 "'program ABC'를 녹화해"라고 발화하고, 'program ABC'에 관련된 복수의 콘텐츠가 디스플레이장치(100)에 표시된 경우를 고려한다. 종래의 경우에는 사용자가 어느 하나를 선택하는 발화를 수행하였을 때에는 정상적으로 어느 한 콘텐츠가 선택됨으로써 녹화가 수행되는 것에 비해, 발화가 아닌 별도의 비발화 입력 방식으로 어느 하나를 선택하였을 때에는 콘텐츠의 선택만이 수행될 뿐 녹화가 수행되지 않는다. 이는, 사용자에 의한 비발화 입력에 의한 정보가 대화형 서버(300)에 전송되지 않으므로, 대화형 서버(300)가 발화 이후의 동작을 취할 수 없기 때문이다.
- [0065] 이러한 점을 고려하여, 본 실시예에 따르면 다음과 같은 방법이 제안된다.
- [0066] 사용자의 입력에 따라서, 디스플레이장치(100)의 대화형 입력 인터페이스(141)에 의해 발생하는 이벤트를 제1이벤트로, 비대화형 입력 인터페이스(142)에 의해 발생하는 이벤트를 제2이벤트라고 편의상 지칭한다. 대화형 서버(300)는 디스플레이장치(100)로부터 제1이벤트가 발생하면 제1이벤트에 대한 처리를 수행하며 제1이벤트의 이력을 저장하고, 이후 제2이벤트가 발생하면 기 저장된 제1이벤트의 처리 이력에 기초하여 제2이벤트의 처리를 수행한다. 여기서, 디스플레이장치(100)는 제2이벤트의 발생 시에, 음성 명령을 인식하게 마련된 대화형 서버(300)가 제2이벤트를 인식할 수 있도록, 제2이벤트를 음성 명령의 형태로 변환하여 대화형 서버(300)에 전송한다.
- [0067] 대화형 서버(300)는 소정 이벤트에 대한 액션을 결정할 때에 이전에 발생했던 이벤트들의 이력을 참조할 수 있다. 대화형 서버(300)는 디스플레이장치(100)의 대화형 입력 인터페이스(141) 및 비대화형 입력 인터페이스(142) 각각에 의해 발생하는 이벤트들을 개별적으로 관리하는 것이 아닌 통합적으로 관리함으로써, 새로 발생한 이벤트에 대응하는 액션을 결정할 때에 대화형 입력 인터페이스(141) 및 비대화형 입력 인터페이스(142)의 여부에 무관하게 이전에 발생했던 이벤트를 참조할 수 있다.
- [0068] 예를 들어 사용자의 최초 발화에 대응하는 검색 결과가 복수 개라면, 대화형 서버(300)는 이 복수의 검색 결과를 사용자가 선택 가능하게 제공한다. 이 때, 사용자가 어느 하나의 검색 결과를 선택하는 지시를 비발화 방식으로 입력하더라도, 대화형 서버(300)는 이벤트 입력의 통합 관리를 통해 신규 이벤트가 이전 이벤트와 입력 방식이 다르더라도 양자간을 상호 관련시킬 수 있다. 이로써, 사용자가 원하는 검색 결과가 도출되어 사용자에게 제공될 수 있다.
- [0069] 도 3 및 도 4는 대화형 서버(300)가 이벤트에 대응하는 검색 결과를 디스플레이장치(100)에 제공하는 방법을 나타내는 플로우차트이다.
- [0070] 도 3에 도시된 바와 같이, S100 단계에서, 대화형 서버(300)는 디스플레이장치(100)로부터 발화에 의한 음성 명령인 제1이벤트를 수신한다.
- [0071] S110 단계에서, 대화형 서버(300)는 제1이벤트를 분석하여 검색 키워드를 도출한다.
- [0072] S120 단계에서, 대화형 서버(300)는 도출된 검색 키워드로 검색을 수행하고, 제1이벤트의 발생 및 처리 이력을 저장한다.
- [0073] S130 단계에서, 대화형 서버(300)는 검색 결과가 복수 개인지 판단한다.
- [0074] 검색 결과가 복수 개가 아니라고 판단하면, S140 단계에서, 대화형 서버(300)는 검색 결과를 디스플레이장치(100)에 전송하고 종료한다.
- [0075] 반면, 검색 결과가 복수 개라고 판단하면, S150 단계에서, 대화형 서버(300)는 복수의 검색 결과를 디스플레이

장치(100)에 전송하고, 복수의 검색 결과 중 어느 하나를 선택할 것을 디스플레이장치(100)에 요청한다.

- [0076] 도 4에 도시된 바와 같이, S160 단계에서, 대화형 서버(300)는 디스플레이장치(100)로부터 선택 지시를 포함하는 제2이벤트를 수신한다. 여기서, 제2이벤트는 디스플레이장치(100)에서 비발화 방식의 사용자 입력에 의해 발생한 것이다.
- [0077] S170 단계에서, 대화형 서버(300)는 기 저장된 제1이벤트의 발생 및 처리 이력을 참조하여 제1이벤트 및 제2이벤트의 관련성, 즉 제2이벤트가 제1이벤트에 종속적으로 처리되어야 할 것인지를 판단한다. 예를 들면, 대화형 서버(300)는 제2이벤트가 제1이벤트에 의해 검색된 복수의 검색 결과에 대한 선택 지시 명령인지를 판단한다.
- [0078] S180 단계에서 제2이벤트가 제1이벤트와 관련성이 없다고 판단되면, 대화형 서버(300)는 S190 단계에서 제1이벤트와 독립적으로 제2이벤트를 분석 처리한다.
- [0079] 반면, S180 단계에서 제2이벤트가 제1이벤트에 의해 검색된 복수의 검색 결과에 대한 선택 지시라고 판단되면, 대화형 서버(300)는 S200 단계에서 제2이벤트에 따라서 검색 결과를 선택한다. 이 과정에서, 필요 여부에 따라서 대화형 서버(300)는 추가적인 검색을 수행할 수 있다.
- [0080] S200 단계에서, 대화형 서버(300)는 제1이벤트 및 제2이벤트에 따른 최종 검색 결과를 디스플레이장치(100)에 전송한다.
- [0081] 이하, 대화형 서버(300)의 구체적인 내부 구성에 관해 도 5를 참조하여 설명한다.
- [0082] 도 5는 대화형 서버(300)의 서버측 프로세서(320)를 기능별로 세분화시킨 구성 블록도이다.
- [0083] 도 5에 도시된 바와 같이, 대화형 서버(300)의 서버측 프로세서(320)는 서버측 통신 인터페이스(310)를 통해 수신되는 음성 명령에 대한 분석을 수행하는 대화 엔진(dialog engine)(321)과, 대화 엔진(321)으로부터 전달받은 검색 키워드에 기초하여 검색을 수행하는 콘텐츠 매니저(322)(content manager)를 포함한다.
- [0084] 대화 엔진(321)은 대화 엔진 매니저(dialog engine manager)(323)와, 대화 에이전트(dialog agent)(324)와, 액션 에이전트(action agent)(325)와, 데이터베이스(326)를 포함한다. 서버측 프로세서(320) 또는 대화 엔진(321)에 대한 이러한 구성요소들의 구분은 기능별로 편의상 적용된 것으로서, 서버측 프로세서(320) 또는 대화 엔진(321)이 반드시 이러한 구분 형태를 가져야 한다는 것을 의미하지는 않는다.
- [0085] 또한, 상기한 구성요소들은 하드웨어적인 구성일 수 있지만, 소프트웨어적인 구성일 수도 있으며, 또는 하드웨어 및 소프트웨어들이 혼용되어 적용될 수도 있다.
- [0086] 대화 엔진 매니저(323)는 대화 엔진(321)에 전달된 음성 명령 형식의 이벤트를 대화 에이전트(324)에 전달한다. 대화 에이전트(324)는 해당 이벤트의 내용을 분석하여, 분석 결과를 구조화된 데이터로 생성하여 대화 엔진 매니저(323)에게 반환한다.
- [0087] 여기서, 대화 에이전트(324)에 의해 생성된 구조화된 데이터를, 대화 프레임(dialog frame)이라고 이하 지칭한다. 대화 프레임은 음성 명령 이벤트의 내용을 포함하되, 대화 엔진 매니저(323), 액션 에이전트(325), 콘텐츠 매니저(322) 등 타 구성요소가 해독 및 참조할 수 있도록 음성 명령을 카테고리 별로 구조화시킨 데이터이다. 카테고리 및 구조화의 형태는 서버측 프로세서(320)의 설계 방식에 따라서 다양하게 결정될 수 있다.
- [0088] 대화 프레임은 음성 명령에서 사용자의 목적과, 그 목적을 달성하기 위해 필요한 속성과, 그 속성의 값에 관한 정보를 포함한다. 대화 프레임에 관한 구체적인 예시는 후술한다.
- [0089] 대화 엔진 매니저(323)는 대화 에이전트(324)로부터 전달받은 대화 프레임을 액션 에이전트(325)에 전달한다. 액션 에이전트(325)는 대화 엔진 매니저(323)로부터 전달받은 대화 프레임에 맞는 액션을 판단하고, 판단 결과를 대화 엔진 매니저(323)에게 반환한다. 즉, 액션 에이전트(325)는 이벤트에 대응하는 액션을 판단한다.
- [0090] 액션 에이전트(325)는 판단 과정에서 정보 검색이 필요한 경우에, 대화 프레임에서 추출한 검색용 키워드를 콘텐츠 매니저(322)에 전달하여 정보 검색을 요청할 수 있다. 콘텐츠 매니저(322)는 전달받은 키워드를 가지고 대화형 서버(300)의 내부 또는 외부에 있는 다양한 종류의 검색 데이터베이스 상에서 검색을 수행하고, 검색 결과를 액션 에이전트(325)에 반환한다.
- [0091] 대화 엔진 매니저(323)는 액션 에이전트(325)로부터 이벤트에 대응하는 액션 또는 검색의 결과를 수신하면, 수신된 결과를 서버측 통신 인터페이스(310)에 전달함으로써 디스플레이장치(100)에 전송되도록 한다.
- [0092] 데이터베이스(326)는 이상과 같은 과정에서 대화 엔진 매니저(323), 대화 에이전트(324), 액션 에이전트(325)에

의해 참조될 수 있는 제반 정보를 포함한다. 예를 들면, 데이터베이스(326)는 대화 에이전트(324)가 대화 프레임 생성하기 위한 컨텍스트(context) 정보, 이벤트 이력 정보, 대화 모델 정보와, 액션 에이전트(325)가 이벤트에 대응하는 액션을 판단하기 위한 액션 모델 정보 등을 포함할 수 있다.

- [0093] 이상의 구조 하에서, 대화형 서버(300)가 디스플레이장치(100)로부터 수신되는 음성 명령에 대응하는 결과를 디스플레이장치(100)에 반환하는 경우의 구체적인 동작 예시에 관해 설명한다.
- [0094] 도 6 및 도 7은 디스플레이장치(100) 및 대화형 서버(300)와, 대화형 서버(300) 내의 각 구성요소들 사이에서 수행되는 데이터의 전송 관계를 나타내는 예시도이다.
- [0095] 도 6에 도시된 바와 같이, 사용자가 "이번 주의 'program ABC'의 출연자는 누구인가?"라는 내용의 발화를 수행하면, 디스플레이장치(100)는 해당 발화에 대응하는 음성 명령을 취득한다. 이 발화에 대응하는 음성 명령을 제1이벤트라고 지칭한다. 디스플레이장치(100)는 제1이벤트를 대화형 서버(300)에 전송한다(410).
- [0096] 대화 엔진 매니저(323)는 서버측 통신 인터페이스(310)에 수신된 제1이벤트를 대화 에이전트(324)에 전달한다(420).
- [0097] 대화 에이전트(324)는 제1이벤트를 분석함으로써, 제1이벤트에 대응하는 대화 프레임을 작성한다. 제1이벤트가 "이번 주의 'program ABC'의 출연자는 누구인가?"의 내용일 때, 대화 프레임은 다음과 같이 표현될 수 있다.
- [0098] (1) 목적: 방송 프로그램의 출연자 문의
- [0099] (2) 제목: program ABC
- [0100] (3) 시간: 이번 주
- [0101] 위 대화 프레임에서, "목적", "제목", "시간"은 대화 프레임의 조건 카테고리이며, 각 조건 카테고리 별로 "방송 프로그램의 출연자 문의", "program ABC", "이번 주"의 속성값을 포함한다. 대화 에이전트(324)는 제1이벤트의 내용을 분석함으로써, 제1이벤트가 포함하는 각 키워드들이 포함되는 조건의 카테고리들과 각 카테고리 별 속성값을 도출할 수 있다.
- [0102] 대화 에이전트(324)로부터 이와 같이 생성된 대화 프레임을 수신하면(430), 대화 엔진 매니저(323)는 대화 프레임을 액션 에이전트(325)에 전달한다(440).
- [0103] 액션 에이전트(325)는 전달받은 대화 프레임에 기초하여, 다음과 같은 키워드를 작성하여 콘텐츠 매니저(322)에 전달한다(450).
- [0104] (1) 제목: program ABC
- [0105] (2) 시간: 2013/06/03~2013/06/09
- [0106] 시간 조건 카테고리의 속성값의 경우, 검색을 위해서는 구체적인 날짜가 명시되어야 한다. 이에, 액션 에이전트(325)는 대화 프레임의 시간 카테고리의 "이번 주"의 속성값에 기초하여, 현재 시간을 기준으로 이번 주의 구체적인 날짜 및 시간 구간을 판단한다.
- [0107] 콘텐츠 매니저(322)는 액션 에이전트(325)로부터 전달받은 키워드로, 다양한 콘텐츠 및 메타데이터의 검색 서비스 상에서 검색을 수행한다. 검색 대상은 콘텐츠 매니저(322)가 액세스할 수 있는 제반 데이터베이스, 검색 서비스, 콘텐츠 서비스가 가능하다.
- [0108] 콘텐츠 매니저(322)는 검색을 통해 나온 검색 결과를 대화 에이전트(324)에 전달한다(460). 예를 들면, 검색 결과는 다음과 같이 복수 개를 포함할 수 있다.
- [0109] (1) program ABC / 2013.06.04 17:00 / 채널 4
- [0110] (2) program ABC / 2013.06.08 10:00 / 채널 20
- [0111] (3) program ABC / 2013.06.09 20:00 / 채널 35
- [0112] 각 검색 결과는 각기 프로그램 제목, 방송시간, 방송채널의 정보를 포함하며, 또한 위에 표시하지는 않았지만 출연자 정보를 더 포함한다. 위 검색 결과에 따르면, "이번 주"에 방송되는 "program ABC"는 세 가지의 경우의 수가 있다.
- [0113] 액션 에이전트(325)는 대화 엔진 매니저(323)에게 검색 결과를 전달한다(470). 이 과정에서, 액션 에이전트

(325)는 검색 결과가 복수이므로 디스플레이장치(100)가 복수의 검색 결과를 선택 가능하게 표시하도록 하는 명령어를 추가한다.

- [0114] 대화 엔진 매니저(323)는 검색 결과 및 명령어를 디스플레이장치(100)에 전송한다(480).
- [0115] 도 7에 도시된 바와 같이, 디스플레이장치(100)는 대화 엔진 매니저(323)로부터 수신된 명령어에 따라서, 복수의 검색 결과를 사용자가 선택 가능하게 표시하고, 표시된 검색 결과 중 어느 하나를 선택할 것을 디스플레이장치(100)의 시스템 발화에 의해 사용자에게 알린다.
- [0116] 이에, 사용자는 리모트 컨트롤러(미도시)의 버튼 조작 또는 제스처와 같은 비발화 방식에 의해, 예를 들어 첫 번째 검색결과와 선택 지시를 입력한다. 이러한 입력을 제2이벤트라고 지칭한다.
- [0117] 디스플레이장치(100)는 제2이벤트에 대응하여, 제2이벤트의 내용을 포함하는 가공의 음성 명령을 생성한다. 여기서, 디스플레이장치(100)는 발화에 의한 음성 명령 형태의 이벤트를 인식 가능하게 마련된 대화형 서버(300)가 제2이벤트를 인식할 수 있도록, 제1이벤트와 유사하게 제2이벤트를 음성 명령의 형태로 변환하여 대화형 서버(300)에 전송한다. 제2이벤트는 예를 들면 "1번"이라는 내용을 포함할 수 있다.
- [0118] 디스플레이장치(100)는 제2이벤트를 대화 엔진 매니저(323)에 전송한다(510).
- [0119] 대화 엔진 매니저(323)는 제2이벤트를 대화 에이전트(324)에 전달한다(520).
- [0120] 제2이벤트가 단지 "1번"이라는 내용만을 포함하고 있는 바, 제2이벤트의 내용만으로 대화 에이전트(324)가 해당 제2이벤트를 처리하는 것은 곤란하다. 이에, 대화 에이전트(324)는 이전 수행된 이벤트 처리 이력로부터 제1이벤트에 관한 처리상태 정보를 호출하고, 호출된 정보에 기초하여 다음과 같은 대화 프레임을 생성한다.
- [0121] (1) 목적: 방송 프로그램의 출연자 문의
- [0122] (2) 순서: 1
- [0123] 본 대화 프레임에서 목적 카테고리 및 그 속성값은 제2이벤트에 포함되어 있는 내용은 아니지만, 대화 에이전트(324)는 제2이벤트와 관련된 제1이벤트를 참조함으로써 위 내용의 대화 프레임을 생성할 수 있다.
- [0124] 대화 에이전트(324)는 생성한 대화 프레임을 대화 엔진 매니저(323)에게 전달하며(530), 대화 엔진 매니저(323)는 대화 프레임을 액션 에이전트(325)에 전달한다(540).
- [0125] 액션 에이전트(325)는 대화 프레임 내의 순서 조건 카테고리의 속성값에 기초하여, 앞선 제1이벤트에 의해 검색된 세 검색 결과 중에서 첫 번째 검색 결과를 선택한다. 액션 에이전트(325)는 "program ABC / 2013.06.04 17:00 / 채널 4"의 선택 결과를 대화 엔진 매니저(323)에게 전달한다(550).
- [0126] 대화 엔진 매니저(323)는 전달받은 선택 결과를 디스플레이장치(100)에 전달한다(560). 이 때, 대화 엔진 매니저(323)는 선택 결과를 참조하여 대화 에이전트(324)에 의해 작성된 시스템 발화 명령어를 추가적으로 디스플레이장치(100)에 전달할 수 있다.
- [0127] 디스플레이장치(100)는 전달받은 명령어에 따라서, 채널 4에서 2013.06.04 17:00에 방송되는 "program ABC"의 출연자를 시스템 발화를 통해 출력한다.
- [0128] 이상 실시예에 따르면, 디스플레이장치(100) 및 대화형 서버(300)에 의해 구현되는 시스템에서 사용자가 발화를 통해 입력한 목표 및 조건을 만족하는 결과물이 복수인 경우에, 사용자가 리모트 컨트롤러나 제스처 등의 비발화 방식의 입력을 혼용하더라도, 사용자의 원래 의도에 맞는 결과를 제공함으로써 사용자의 편의성을 향상시킬 수 있다.
- [0129] 한편, 위 실시예에서는 대화형 서버(300)가 디스플레이장치(100)의 사용자의 발화 방식 입력 및 비발화 방식의 입력 각각에 의한 이벤트의 이력을 저장하고 있는 경우에 관해 설명하였으나, 본 발명의 사상이 이에 한정되지 않는다. 예를 들면, 디스플레이장치(100)가 이벤트의 이력을 저장하는 구성도 가능한 바, 이에 관하여 이하 도 8 및 도 9를 참조하여 설명한다.
- [0130] 도 8 및 도 9는 본 발명의 제2실시예에 따른 디스플레이장치(100)의 제어방법을 나타내는 플로우차트이다.
- [0131] 도 8에 도시된 바와 같이, S300 단계에서, 디스플레이장치(100)는 사용자의 발화를 수신한다. S310 단계에서, 디스플레이장치(100)는 사용자의 발화에 대응하는 음성 명령을 대화형 서버(300)에 전송한다.
- [0132] S320 단계에서, 디스플레이장치(100)는 대화형 서버(300)로부터 수신된 음성 명령의 대응 결과가 복수 개인지

여부를 판단한다.

- [0133] 대화형 서버(300)로부터 수신된 음성 명령의 대응 결과가 하나이면, S330 단계에서, 디스플레이장치(100)는 해당 결과를 표시한다.
- [0134] 반면, 대화형 서버(300)로부터 수신된 음성 명령의 대응 결과가 복수 개이면, S340 단계에서, 디스플레이장치(100)는 복수의 대응 결과를 사용자가 선택 가능하게 표시한다.
- [0135] S350 단계에서, 디스플레이장치(100)는 사용자로부터 복수의 대응 결과에 대한 선택 입력을 수신한다. 이상의 과정이 진행되는 동안에, 디스플레이장치(100)는 발화에 의한 음성 명령의 이벤트의 발생 및 처리 과정의 이력을 저장한다.
- [0136] 도 9에 도시된 바와 같이, S360 단계에서, 디스플레이장치(100)는 수신된 선택 입력이 발화 방식인지 아니면 비발화 방식인지 여부를 판단한다.
- [0137] 이 과정에서 디스플레이장치(100)는 기 저장된 이전의 이벤트 이력을 참조하여, S350 단계에서 수신된 선택 입력이 S300 단계에서 발생한 발화와 관련성이 있는지 판단한다. 디스플레이장치(100)는 판단 결과 양자간에 관련성이 있다고 판단되는 경우에 S360 단계를 수행한다.
- [0138] 수신된 선택 입력이 발화 방식이면, S370 단계에서, 디스플레이장치(100)는 선택 입력에 대응하는 음성 명령을 대화형 서버(300)에 전송하고 S400 단계로 이행한다.
- [0139] 반면 수신된 선택 입력이 비발화 방식이면, S380 단계에서, 디스플레이장치(100)는 비발화 방식의 선택 입력의 내용에 대응하는 가공의 음성 명령을 생성한다. S390 단계에서, 디스플레이장치(100)는 이와 같이 생성된 가공의 음성 명령을 대화형 서버(300)에 전송한다.
- [0140] S400 단계에서, 디스플레이장치(100)는 전송한 선택 입력에 대응하는 결과를 대화형 서버(300)로부터 수신하여 표시한다.
- [0141] 한편, 이상 실시예에서는 디스플레이장치(100)에서 발생한 이벤트에 대응하는 결과를 대화형 서버(300)가 분석 및 검색하는 경우에 관해 설명하였다. 그러나, 본 발명의 사상이 이에 한정되지 않는 바, 제1실시예에서의 STT 서버(200) 및 대화형 서버(300) 중 적어도 어느 하나의 기능을 디스플레이장치(100)가 수행하는 실시예도 가능하다.
- [0142] 도 10은 본 발명의 제3실시예에 따른 디스플레이장치(600)의 프로세서(620)의 구성 블록도이다.
- [0143] 도 10에 도시된 바와 같이, 본 실시예에 따른 디스플레이장치(600)는 통신 인터페이스(610)와 프로세서(620)를 포함한다. 디스플레이장치(600)가 포함하는 그 외의 구성요소들에 관해서는 제1실시예를 응용할 수 있는 바, 자세한 설명을 생략한다.
- [0144] 프로세서(620)는 디스플레이장치(600)에 입력되는 발화를 처리하는 음성처리엔진(621)과, 발화에 대응하는 음성 명령에 대한 분석을 수행하는 대화 엔진(622)과, 대화 엔진(622)으로부터 전달받은 검색 키워드에 기초하여 검색을 수행하는 콘텐츠 매니저(623)를 포함한다. 또한, 대화 엔진(622)은 대화 엔진 매니저(624)와, 대화 에이전트(625)와, 액션 에이전트(626)와, 데이터베이스(627)를 포함한다. 본 도면 및 실시예에서는 발명의 사상에 직접적으로 관련이 있는 구성에 관해서만 설명하고 있으며, 그 이외의 구성에 관해서는 도시 및 설명을 생략하였음을 밝힌다.
- [0145] 음성처리엔진(621)은 사용자의 발화가 입력되면 해당 발화의 음성신호를 STT 서버(200)에 전송하고, 해당 발화에 대응하는 음성 명령을 STT 서버(200)로부터 수신하여 대화 엔진 매니저(624)에게 전달한다.
- [0146] 이하 대화 엔진 매니저(624), 대화 에이전트(625), 액션 에이전트(626) 및 콘텐츠 매니저(623)에 의한 구체적인 동작은 앞선 실시예에서 설명한 동일 명칭의 구성요소와 실질적으로 동일하므로, 앞선 실시예의 설명을 응용할 수 있는 바, 자세한 설명을 생략한다.
- [0147] 또는, 디스플레이장치(600)는 음성처리엔진(621)이 STT 서버(200)의 기능, 즉 발화를 음성 명령으로 변환하는 기능을 수행할 수도 있다. 이 경우, 음성처리엔진(621)은 사용자의 발화가 입력되면 해당 발화를 음성 명령으로 변환하여 대화 엔진 매니저(624)에게 전달한다.
- [0148] 상기한 실시예는 예시적인 것에 불과한 것으로, 당해 기술 분야의 통상의 지식을 가진 자라면 다양한 변형 및 균등한 타 실시예가 가능하다. 따라서, 본 발명의 진정한 기술적 보호범위는 하기의 특허청구범위에 기재된 발

명의 기술적 사상에 의해 정해져야 할 것이다.

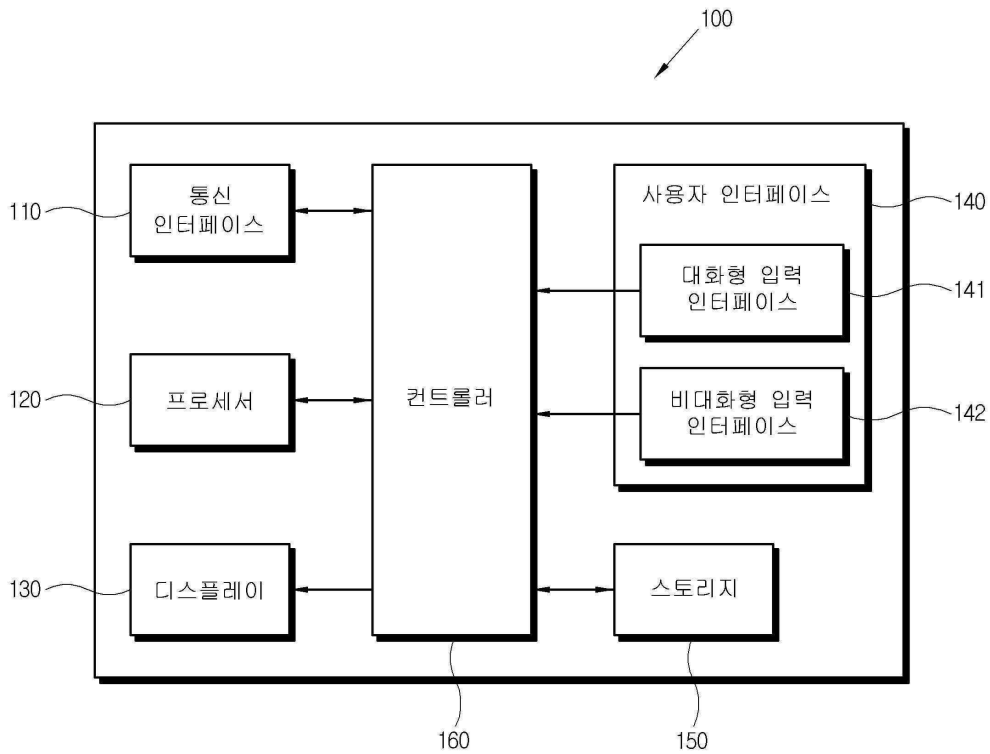
부호의 설명

[0149]

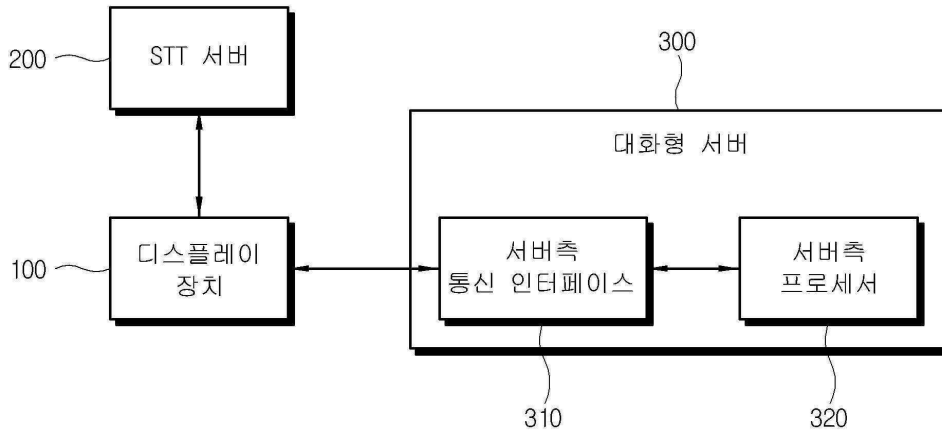
- 100 : 디스플레이장치
- 110 : 통신 인터페이스
- 120 : 프로세서
- 130 : 디스플레이
- 140 : 사용자 인터페이스
- 141 : 대화형 입력 인터페이스
- 142 : 비대화형 입력 인터페이스
- 150 : 스토리지
- 160 : 컨트롤러
- 200 : STT 서버
- 300 : 대화형 서버
- 310 : 서버측 통신 인터페이스
- 320 : 서버측 프로세서

도면

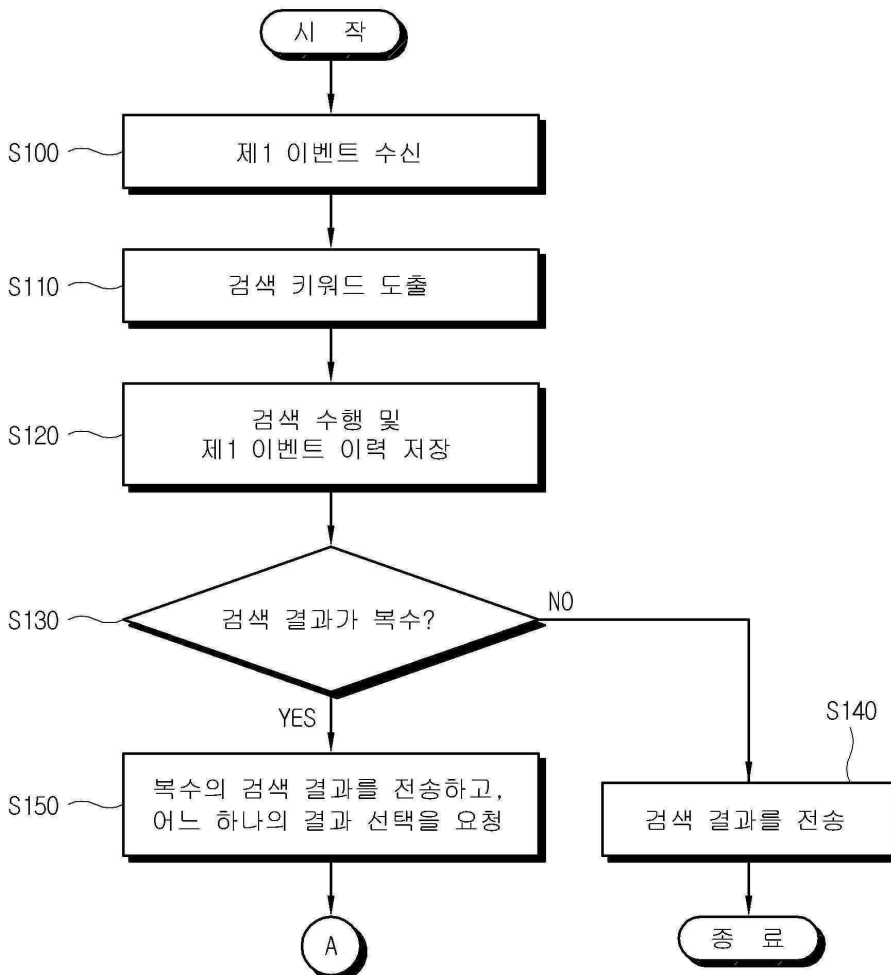
도면1



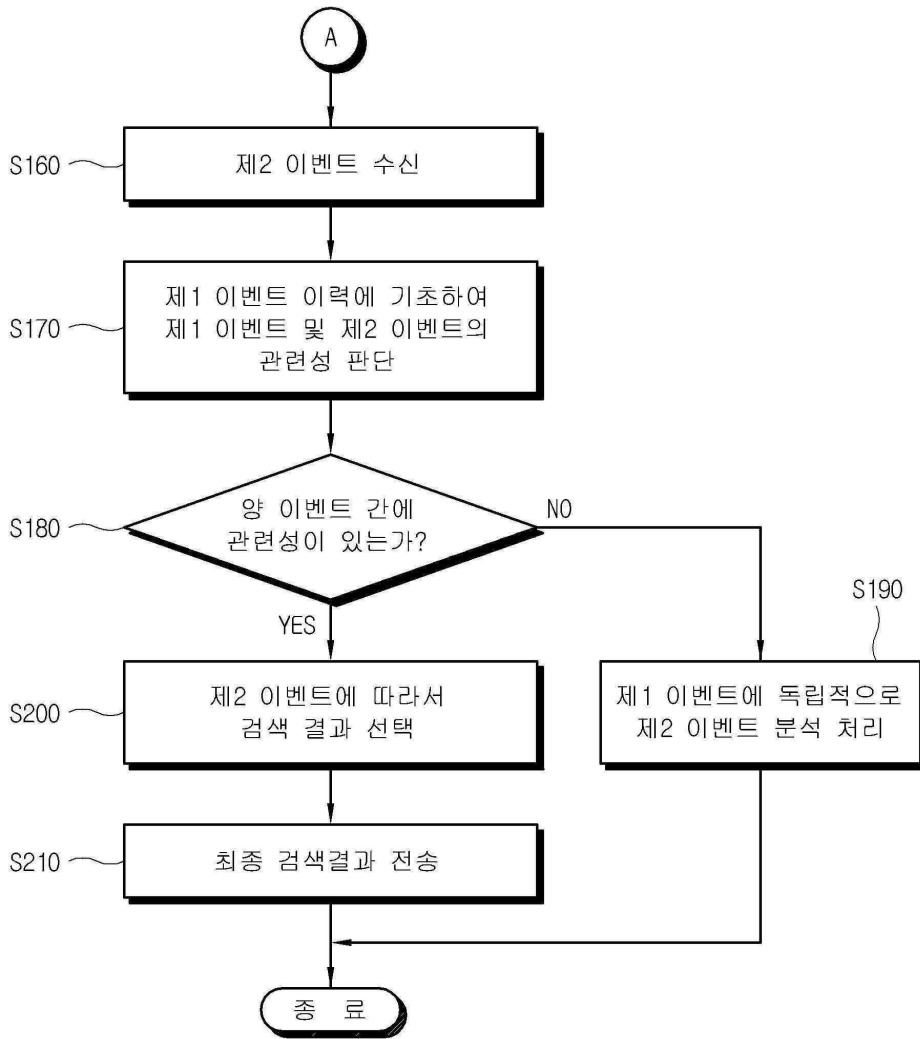
도면2



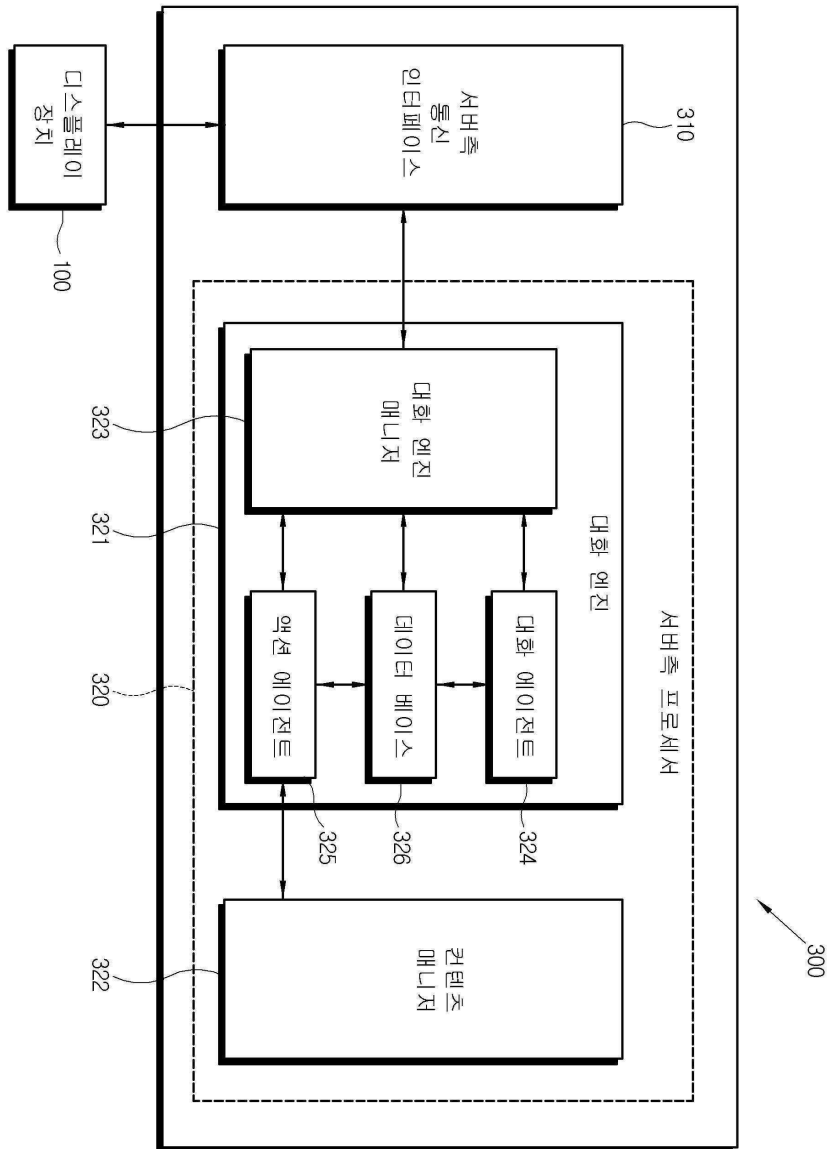
도면3



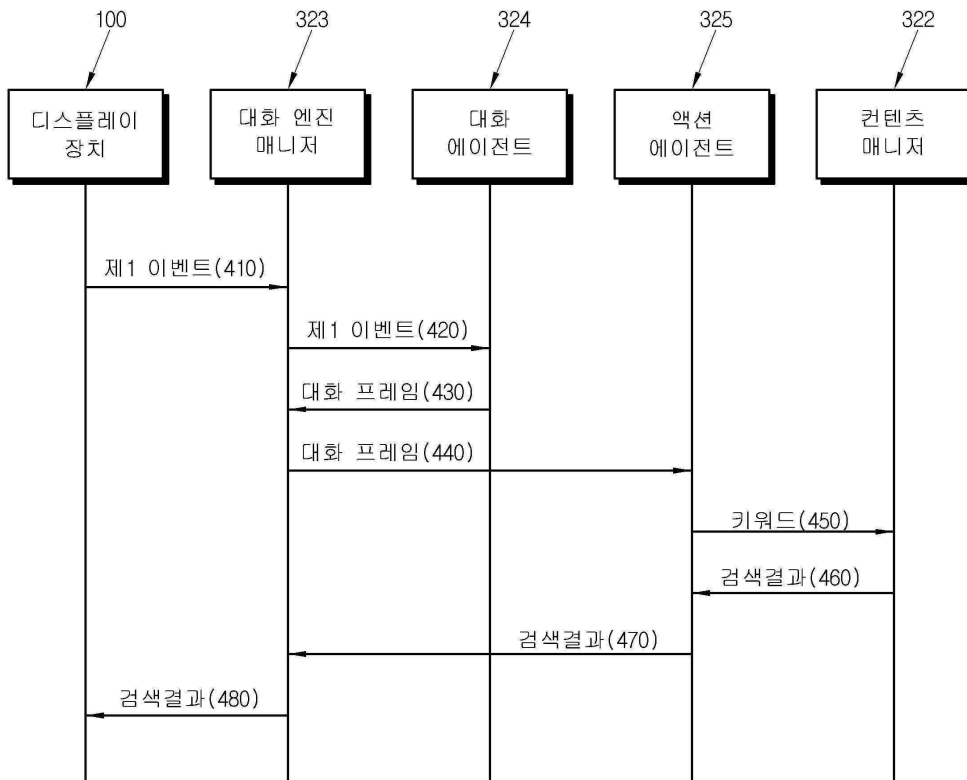
도면4



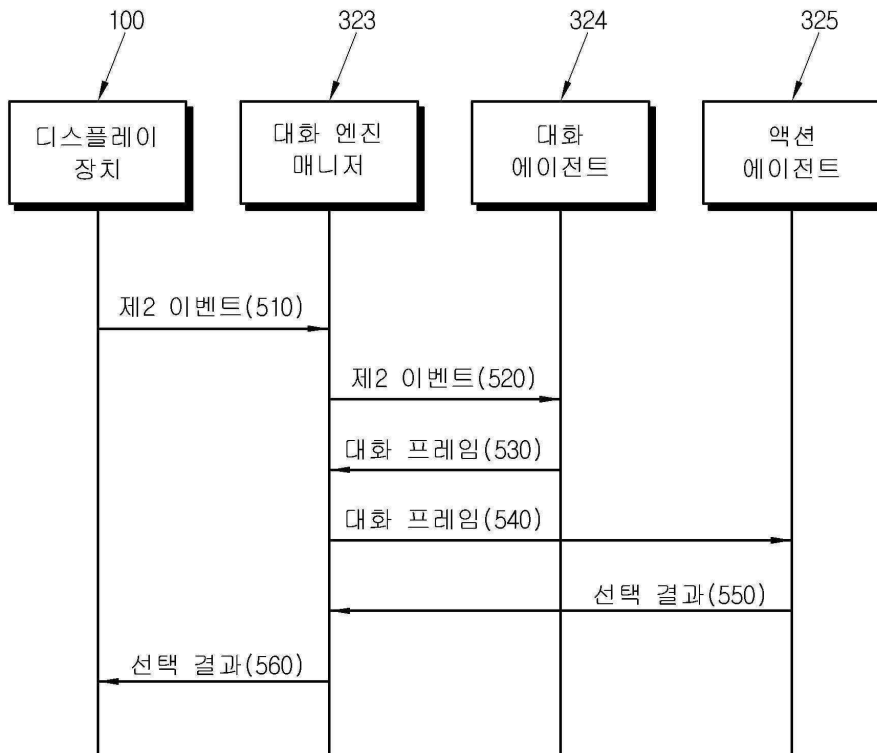
도면5



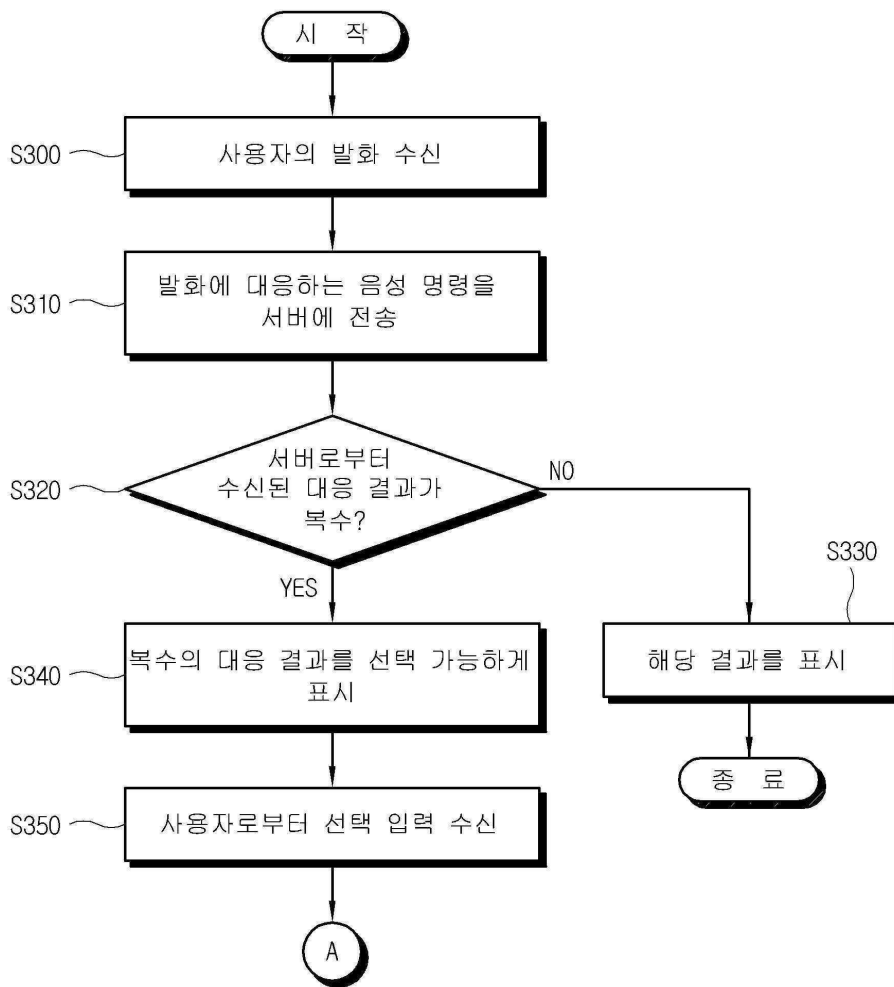
도면6



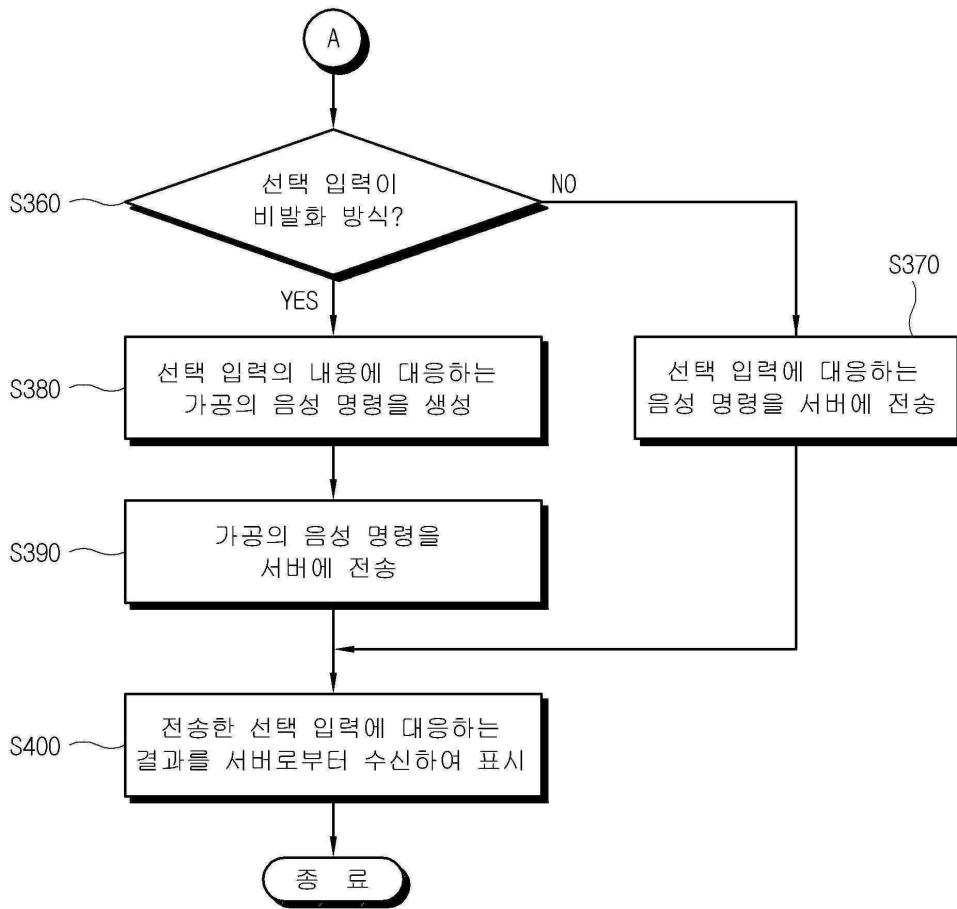
도면7



도면8



도면9



도면10

