



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2012년02월06일  
(11) 등록번호 10-1109293  
(24) 등록일자 2012년01월17일

(51) Int. Cl.

G06F 15/16 (2006.01) G06F 17/00 (2006.01)

(21) 출원번호 10-2004-0084562

(22) 출원일자 2004년10월21일

심사청구일자 2009년10월14일

(65) 공개번호 10-2005-0045818

(43) 공개일자 2005년05월17일

(30) 우선권주장

10/705,155 2003년11월11일 미국(US)

(56) 선행기술조사문헌

KR 2001-0008284 A

(73) 특허권자

마이크로소프트 코포레이션

미국 워싱턴주 (우편번호 : 98052) 레드몬드 원  
마이크로소프트 웨이

(72) 발명자

혼시아오-우웬

미국 98052 워싱턴주 레드몬드 원 마이크로소프트  
웨이 마이크로소프트 코포레이션 내

왕구안산

미국 98052 워싱턴주 레드몬드 원 마이크로소프트  
웨이 마이크로소프트 코포레이션 내

(74) 대리인

제일특허법인

전체 청구항 수 : 총 20 항

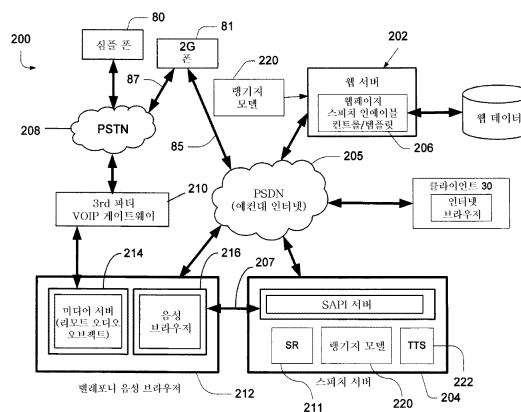
심사관 : 퇴-안철용

(54) 순차 멀티모드 입력

(57) 요약

2G 모바일폰과 클라이언트/서버 아키텍처의 인터랙트 방법이 개시된다. 상기 2G 폰은 데이터 송신용의 데이터 채널과 스피치 송신용의 음성 채널을 구비한다. 본 방법은 상기 데이터 채널을 통해 애플리케이션에 따라서 웹 서버로부터 웹페이지를 수신하여 상기 웹페이지를 상기 2G 폰에 렌더링하는 단계를 포함한다. 상기 웹페이지 상의 적어도 하나의 데이터 필드에 대응하는 사용자로부터 스피치가 수신된다. 상기 음성 채널을 통해 상기 2G 폰으로부터 텔레포니 서버로의 통화(call)가 확립된다. 상기 텔레포니 서버는 상기 2G 폰으로부터 원격지에 위치하며 스피치를 처리하도록 구성된다. 상기 텔레포니 서버는 상기 2G 폰에 제공된 상기 웹페이지에 대응하는 웹 서버로부터 스피치 인에이블(speech-enabled) 웹페이지를 획득한다. 상기 스피치는 상기 2G 폰으로부터 상기 텔레포니 서버에 송신된다. 상기 스피치는 상기 스피치 인에이블 웹페이지에 따라서 처리되어 텍스트 데이터가 획득된다. 상기 텍스트 데이터는 상기 웹 서버에 송신된다. 상기 2G 폰은 상기 데이터 채널을 통해 새로운 웹페이지를 획득하고 상기 텍스트 데이터를 갖는 상기 새로운 웹페이지를 렌더링한다.

대표도



## 특허청구의 범위

### 청구항 1

데이터 송신용의 데이터 채널과 스피치 송신용의 음성 채널을 갖는 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법으로서,

상기 데이터 채널을 통해 애플리케이션에 따라서 웹서버로부터 적어도 하나의 데이터 필드를 포함하는 웹페이지를 수신하고, 상기 2G 모바일폰 상에서 상기 웹페이지를 렌더링하는 단계;

상기 웹페이지 상의 상기 적어도 하나의 데이터 필드에 대응하는 스피치를 사용자로부터 수신하는 단계;

상기 음성 채널을 통해 상기 2G 모바일폰으로부터 텔레포니 서버로의 통화(call)를 확립하는 단계 - 상기 텔레포니 서버는 상기 2G 모바일폰으로부터 원격지에 위치하며 스피치를 처리하도록 구성됨 -;

상기 텔레포니 서버에 의해 상기 웹서버로부터, 상기 2G 모바일폰에 제공된 상기 웹페이지에 대응하는 스피치 인에이블(speech-enabled) 웹페이지를 획득하는 단계;

상기 2G 모바일폰으로부터 상기 텔레포니 서버로 상기 스피치를 송신하는 단계;

상기 스피치에 따른 텍스트 데이터를 획득하기 위해 상기 텔레포니 서버에 의해 상기 스피치 인에이블 웹페이지에 따라 상기 스피치를 처리하는 단계 - 상기 텍스트 데이터는 상기 적어도 하나의 데이터 필드에 대응함 -;

상기 텔레포니 서버로부터 상기 웹서버로 상기 텍스트 데이터를 송신하는 단계; 및

상기 데이터 채널을 통해 상기 2G 모바일폰 상에서 새로운 웹페이지를 획득하고, 대응하는 상기 적어도 하나의 데이터 필드에 상기 텍스트 데이터가 입력된 상기 새로운 웹페이지를 렌더링하는 단계

를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

### 청구항 2

제1항에 있어서,

상기 스피치를 처리하는 단계는, 상기 수신된 스피치를 나타내는 데이터를 상기 텔레포니 서버로부터 원격지에 있는 스피치 서버에 송신하는 단계를 포함하고, 상기 스피치 서버는 상기 수신된 스피치를 나타내는 상기 데이터를 처리하여 상기 텍스트 데이터를 획득하며, 상기 웹서버로 상기 텍스트 데이터를 송신하는 단계는 상기 스피치 서버가 상기 텍스트 데이터를 송신하는 단계를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

### 청구항 3

제1항에 있어서,

상기 음성 채널을 통해 상기 2G 모바일폰으로부터 텔레포니 서버로의 통화를 확립하는 단계는 상기 2G 모바일폰과 연관된 식별자를 획득하는 단계를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

### 청구항 4

제3항에 있어서,

상기 웹서버로부터, 상기 2G 모바일폰에 제공된 상기 웹페이지에 대응하는 스피치 인에이블 웹페이지를 획득하는 단계는 상기 2G 모바일폰과 연관된 상기 식별자를 이용하는 단계를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

### 청구항 5

제4항에 있어서,

상기 식별자를 획득하는 단계는 상기 2G 모바일폰과 연관된 전화번호를 식별하는 단계를 포함하는, 2G 모바일폰

을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 6

제1항에 있어서,

상기 웹서버로부터 상기 새로운 웹페이지를 획득하는 단계에 앞서, 새로운 페이지가 상기 웹서버로부터 입수가 가능하다는 메시지를 상기 2G 모바일폰에 송신하는 단계를 더 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 7

제6항에 있어서,

상기 메시지를 송신하는 단계는 SMS 메시지를 송신하는 단계를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 8

제6항에 있어서,

상기 메시지를 송신하는 단계는 상기 새로운 웹페이지에 대한 어드레스와 관련되는 정보를 송신하는 단계를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 9

제8항에 있어서,

상기 메시지를 송신하는 단계는 URL 링크를 송신하는 단계를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 10

제9항에 있어서,

상기 메시지를 송신하는 단계는 SMS 메시지를 송신하는 단계를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 11

제6항에 있어서,

상기 텔레포니 서버가 상기 새로운 웹페이지를 획득하는 단계에 앞서, 상기 음성 채널을 차단(disconnect)하는 단계를 더 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 12

제2항에 있어서,

상기 원격지에 있는 스피치 서버는 상기 웹서버 상의 상기 애플리케이션의 제공자와 상이한 제3자에 의해 제공 되도록 구성되는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 13

데이터 송신용의 데이터 채널과 스피치 송신용의 음성 채널을 갖는 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법으로서,

상기 데이터 채널을 통해 애플리케이션에 따라서 웹서버로부터 적어도 하나의 데이터 필드를 포함하는 웹페이지를 수신하고, 상기 2G 모바일폰 상에서 상기 웹페이지를 렌더링하는 단계;

상기 웹페이지 상의 상기 적어도 하나의 데이터 필드에 대응하는 스피치를 사용자로부터 수신하는 단계;

상기 음성 채널을 통해 상기 2G 모바일폰으로부터 텔레포니 서버로의 통화를 확립하는 단계 - 상기 텔레포니 서

버는 상기 2G 모바일폰으로부터 원격지에 위치하며 스피치를 처리하도록 구성됨 -;

상기 2G 모바일폰으로부터 상기 텔레포니 서버로 상기 스피치를 송신하는 단계; 및

상기 데이터 채널을 통해 상기 2G 모바일폰 상에서 갱신된 웹페이지를 획득하고, 상기 스피치에 따라 상기 대응하는 적어도 하나의 데이터 필드가 갱신된 상기 갱신된 웹페이지를 렌더링하는 단계

를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 14

제13항에 있어서,

상기 음성 채널을 통해 상기 2G 모바일폰으로부터 텔레포니 서버로의 통화를 확립하는 단계는 상기 2G 모바일폰과 연관된 식별자를 송신하는 단계를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 15

제14항에 있어서,

상기 식별자를 송신하는 단계는 상기 2G 모바일폰과 연관된 전화번호를 식별하는 단계를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 16

제13항에 있어서,

상기 웹서버로부터 상기 갱신된 웹페이지를 획득하는 단계에 앞서, 갱신된 페이지가 상기 웹서버로부터 입수가 가능하다는 메시지를 수신하는 단계를 더 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 17

제16항에 있어서,

상기 메시지를 수신하는 단계는 SMS 메시지를 수신하는 단계를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 18

제16항에 있어서,

상기 메시지를 수신하는 단계는 상기 갱신된 웹페이지에 대한 어드레스와 관련되는 정보를 수신하는 단계를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 19

제18항에 있어서,

상기 메시지를 수신하는 단계는 URL 링크를 수신하는 단계를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

#### 청구항 20

데이터 송신용의 데이터 채널과 스피치 송신용의 음성 채널을 갖는 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법으로서,

2G 모바일폰 상에서 렌더링하기에 적합하도록, 상기 데이터 채널을 통해 애플리케이션에 따라서 웹서버로부터 적어도 하나의 데이터 필드를 포함하는 웹페이지를 송신하는 단계;

상기 웹서버로부터 텔레포니 서버로, 상기 2G 모바일폰에 제공된 상기 웹페이지에 대응하는 스피치 인에이블 웹페이지를 송신하는 단계;

상기 적어도 하나의 데이터 필드에 대한 인식된 스피치를 나타내는 상기 텔레포니 서버로부터의 텍스트 데이터를 수신하는 단계; 및

대응하는 상기 적어도 하나의 데이터 필드에 상기 텍스트 데이터가 입력된 갱신된 웹페이지를 상기 데이터 채널을 통해 상기 2G 모바일폰으로 송신하는 단계

를 포함하는, 2G 모바일폰을 이용하여 클라이언트/서버 아키텍처와 인터랙트하는 방법.

## 명 세 서

### 발명의 상세한 설명

#### 발명의 목적

#### 발명이 속하는 기술 및 그 분야의 종래기술

[0024] 본 발명은 컴퓨터 시스템에서의 정보의 액세스 및 렌더링에 관한 것이다. 보다 구체적으로, 본 발명은 2세대 ("2G") 모바일폰 또는 셀룰러폰에서의 순차 멀티모드 입력에 관한 것이다.

[0025] 휴대폰이나 PIM(personal information managers)과 같은 소형 컴퓨팅 장치가 사람들의 일상 활동에서 점점 더 많이 이용되고 있다. 이들 장치를 동작시키는데 사용되는 마이크로프로세서의 처리 능력이 증대됨에 따라 이들 장치의 기능이 늘어나고 경우에 따라서는 이들이 합쳐지고 있다. 예를 들어, 특히 2G 폰을 포함하여 많은 휴대폰들은 현재 이를 사용하여 인터넷에 액세스 및 브라우즈할 수 있을뿐만 아니라, 어드레스, 전화번호 등의 개인정보를 저장하는데도 이용할 수 있다.

[0026] 이들 컴퓨팅 장치들이 인터넷 브라우징에 이용되고 있다는 점이나 다른 서버/클라이언트 아키텍처에 이용되고 있다는 점에서, 이들 컴퓨팅 장치에 정보를 입력하는 것이 필요하게 된다. 그러나 이들 장치의 휴대를 용이하게 하기 위해서는 이들 장치를 가능한 한 작게 유지해야 하므로, 상기한 컴퓨팅 장치의 하우징에서 이용가능한 표면적이 한정되어 모든 문자를 분리된 버튼으로서 구비하고 있는 통상의 키보드는 통상적으로 불가능하다. 따라서, 인터넷과 같은 클라이언트/서버 아키텍처를 내비게이션하기 위해 이러한 장치의 사용자는 웹페이지의 필요한 필드를 채울 텍스트 정보를 제공하거나 또는 명령을 제공하는 방식으로 한정된 키보드를 조작해야 한다. 이러한 방식의 입력은 이러한 구속하에 동작하는 웹 기반 애플리케이션의 유용성을 제한하며, 이러한 장치를 이용한 인터넷 또는 다른 클라이언트/서버 시스템의 내비게이션은 눈에 띄는 성공을 이루지 못했다.

[0027] 최근에는 음성 포털이 SALT(Speech Application Language Tags) 또는 VoiceXML(voice extensible markup language) 등의 이용을 통해 단지 전화를 이용하여 인터넷 콘텐츠에 액세스할 수 있도록 진화되었다. 이러한 아키텍처에서, 문서 서버(예컨대 웹서버)는 SALT/VoiceXML 번역기를 통해 클라이언트로부터의 요청을 처리하게 된다. 웹서버는 이에 응답하여 SALT/VoiceXML 번역기에 의해 처리되고 사용자의 청취가 가능하도록 렌더링된 SALT/VoiceXML 문서를 생성할 수 있게 된다. 사용자는 음성 인식을 통한 음성 명령을 이용하여 웹을 내비게이션할 수 있게 된다. 특히 웹서버로부터 획득한 정보가 청취가 가능하도록 렌더링되어야 하므로 사용자에게 렌더링되어 리턴되는 경우에는 이러한 인터넷 내비게이션 기술도 제한된다. 또한, 사용자는 인식 결과를 눈으로 확인하지 않고는 인식이 제대로 되었는지를 확인할 수 없다. 인식된 결과를 가청음으로써 확인시켜 줄 수는 있지만 이러한 확인에는 시간이 걸리며 따라서 사용자의 원활하고(streamlined) 효율적인 체험을 방해한다.

#### 발명이 이루고자 하는 기술적 과제

[0028] 따라서 서버/클라이언트 아키텍처에서 특히 2G 폰과 같은 장치에 대한 서버 정보를 액세스하는데 이용하기 위한 아키텍처 및 방법을 개선할 필요가 있다.

#### 발명의 구성 및 작용

[0029] 2G 모바일폰과 클라이언트/서버 아키텍처의 인터랙트 방법이 개시된다. 상기 2G 폰은 데이터 송신용의 데이터 채널과 스피치 송신용의 음성 채널을 구비한다. 본 방법은 상기 데이터 채널을 통해 애플리케이션에 따라서 웹서버로부터 웹페이지를 수신하여 상기 웹페이지를 상기 2G 폰에 렌더링하는 단계를 포함한다. 상기 웹페이지 상의 적어도 하나의 데이터 필드에 대응하는 사용자로부터 스피치가 수신된다. 상기 음성 채널을 통해 상기 2G 폰으로부터 텔레포니 서버로의 통화(call)가 확립된다. 상기 텔레포니 서버는 상기 2G 폰으로부터 원격지에 위치하며 스피치를 처리하도록 구성된다. 상기 텔레포니 서버는 상기 2G 폰에 제공된 상기 웹페이지에 대응하는

웹서버로부터 스피치 인에이블(speech-enabled) 웹페이지를 획득한다. 상기 스피치는 상기 2G 폰으로부터 상기 텔레포니 서버에 송신된다. 상기 스피치는 상기 스피치 인에이블 웹페이지에 따라서 처리되어 텍스트 데이터가 획득된다. 상기 텍스트 데이터는 상기 웹서버에 송신된다. 상기 2G 폰은 상기 데이터 채널을 통해 새로운 웹페이지를 획득하고 상기 텍스트 데이터를 갖는 상기 새로운 웹페이지를 렌더링한다.

[0030] 본 발명의 또다른 태양으로서 상기한 2G 폰의 동작 관점에서 살펴보면, 본 방법은 상기 데이터 채널을 통해 애플리케이션에 따라서 웹서버로부터 웹페이지를 수신하여 상기 웹페이지를 상기 2G 폰에 렌더링하는 단계를 포함한다. 상기 웹페이지 상의 적어도 하나의 데이터 필드에 대응하는 사용자로부터 스피치가 수신된다. 상기 음성 채널을 통해 상기 2G 폰으로부터 텔레포니 서버로의 통화(call)가 확립되며, 상기 텔레포니 서버는 상기 2G 폰으로부터 원격지에 위치하며 스피치를 처리하도록 구성된다. 상기 스피치는 상기 2G 폰으로부터 상기 텔레포니 서버에 송신된다. 상기 데이터 채널을 통해 상기 2G 폰에 새로운 웹페이지가 획득되며 상기 스피치에 따른 텍스트 데이터를 갖는 상기 새로운 웹페이지가 렌더링된다.

[0031] <실시예>

[0032] 본 발명의 일 태양은 스피치 인식이 구현된 멀티모드 입력을 2세대("2G") 폰에 제공하는 방법과 관련된다. 본 명세서에서 2G 폰은 공지된 바와 같이 음성 채널을 통해 음성 통화를 할 수 있으며 별도의 데이터 채널을 통해 디지털 데이터를 발신 및 수신할 수 있는 회로부를 포함하고 있다. 사용자는 폰의 키패드를 이용하여 클라이언트/서버 아키텍처 내의 웹사이트를 내비게이트하고 텍스트 데이터를 발신 및 수신함으로써 정보를 획득할 수 있다. 상기한 데이터는 작은 디스플레이 상에 렌더링된다. 본 발명의 일 태양에 따르면 사용자가 폰 입력의 한 형태로서 스피치를 제공할 수 있도록 함으로써 대응하는 텍스트를 입력하는 번거로운 작업을 회피하고 있다.

[0033] 도 5를 참조하면 본 발명에 채용될 수 있는 웹기반 스피치 인식 아키텍처(200)가 도시되어 있다. 일반적으로 웹서버(202)에 저장된 정보의 액세스는 모바일 장치(30) - 본 명세서에서는 가청신호를 검출하는 마이크뿐만 아니라 표시화면을 구비한 다른 형태의 컴퓨팅 장치를 나타내기도 함 - 를 통해 수행되거나; 심플 폰(80)을 통해 수행되어 눌려진 키에 응답하여 폰(80)에서 발생된 톤(tone)을 통해 또는 청취 가능한 형태로 정보가 요청되며, 또한 웹서버(202)로부터의 정보가 사용자에게 단지 청취가 가능하도록 제공되거나; 또는 2G 폰(81)을 통해 수행되어 정보를 웹서버(202)로부터도 액세스할 수 있으며 상기 정보가 WAP(Wireless Application Protocol)을 통해 송신된 WML 또는 XHTML 페이지 등의 페이지로서 제공되게 된다. 전술한 바와 같이, 제한된 키보드 능력을 감안하면, 본 발명에서 채용하고 있는 상기한 아키텍처(200)는 2G 폰(81)의 유용성을 개량하기 위해 스피치 인식과 함께 사용할 수 있도록 하는 한편, 2G 폰의 비주얼 렌더링 능력을 이용하여 인식결과를 렌더링할 수 있다.

[0034] 보다 중요한 것은, 스피치 인식을 이용하여 장치(30), 심플 폰(80) 또는 2G 폰(81)을 통해 정보가 획득된다는 점에서 상기한 아키텍처(200)는 일원화되어 있지만, 싱글 스피치 서버(204)가 각각의 동작 모드를 서포트할 수 있다는 점이다. 또한, 상기한 아키텍처(200)는 각종의(an extension of) 공지된 마크업 랭기지(예컨대, HTML, XHTML, cHTML, XML, WML 등)를 이용하여 동작한다. 따라서 웹서버(202)에 저장된 정보는 이들 마크업 랭기지에서 제공하는 공지의 GUI 방식을 이용하여 액세스될 수 있다. 각종의 공지된 마크업 랭기지를 이용함으로써 웹서버(202)에서의 프로그래밍(authoring)이 용이하게 되며 또한 음성 인식을 포함하도록 기존의 레거시(legacy) 애플리케이션을 용이하게 수정할 수 있다.

[0035] 웹기반 스피치 인식 아키텍처(200) 및 2G 폰(81)에서의 웹기반 스피치 인식의 구현 방법에 대한 설명에 앞서, 이러한 아키텍처(200)에서 동작하는 다른 컴퓨팅 장치에 대해 전반적으로 설명하기로 한다.

[0036] 도 1을 참조하면, 데이터 관리장치(PIM, PDA 등)의 일례를 30으로 도시하고 있다. 모바일 장치(30)는 하우징(32)을 포함하며 디스플레이(34)를 포함하는 사용자 인터페이스를 구비하며, 사용자 디스플레이는 스타일러스(33)와 연계한 접촉 감응식 표시화면을 사용한다. 스타일러스(33)를 사용하여 지정된 좌표에 있는 디스플레이(34)를 누르거나 접촉하게 되면 필드를 선택하거나 또는 커서의 시작위치를 선택적으로 이동시키거나 또는 커맨드 정보를 제공하게 된다. 다른 방법으로 또한 내비게이션을 위해 장치(30)에 하나 이상의 버튼(35)이 구비될 수도 있다. 또한 회전가능 휠, 롤러 등의 기타 입력 메커니즘이 구비될 수도 있다.

[0037] 도 2를 참조하면, 상기한 모바일 장치(30)를 구비한 기능 컴포넌트를 블록도로서 도시하고 있다. 중앙처리장치(CPU)(50)는 소프트웨어 제어 기능을 구현한다. CPU(50)는 디스플레이(34)에 연결되어 소프트웨어 제어에 따라 생성된 텍스트 및 그래픽 아이콘이 디스플레이(34)에 나타나게 된다. CPU(50)에는 통상적으로 디지털/아날로그 변환기(59)를 구비한 스피커(43)가 연결되어 가청 출력을 제공한다. 사용자에 의해 모바일 장치(30)에 다운로드 또는 입력된 데이터는 상기한 CPU(50)에 양방향으로 연결되어 있는 불활성 판독/기록 랜덤 액세스 메모리 저



장부(54)에 저장된다. RAM(랜덤 액세스 메모리)(54)은 CPU(50)에 의해 실행되는 명령에 대한 활성 저장부를 제공하며 또한 레지스터값과 같은 임시 데이터에 대한 저장부를 제공한다. 설정 옵션을 위한 디폴트값 및 기타 변수는 판독 전용 메모리(ROM)(58)에 저장된다. ROM(58)은 또한 모바일 장치(30)의 기본 기능과 기타 오퍼레이팅 시스템 커널 기능(예컨대, 소프트웨어 컴포넌트의 RAM(54)으로의 로딩)을 제어하는 상기한 장치용의 오퍼레이팅 시스템 소프트웨어를 저장하는데 이용된다. RAM(54)은 또한 애플리케이션 프로그램을 기록하는데 이용되는 PC 상의 하드드라이브 기능과 유사한 방식으로 코드에 대한 저장부로서 동작한다.

[0038] 모바일 장치는 CPU(50)에 연결된 무선 트랜시버(52)를 통해 무선 신호를 송신/수신할 수 있다. 컴퓨터(예컨대 데스크톱 컴퓨터)로부터 직접 또는 원하는 경우 유선 네트워크를 통해 데이터를 다운로드하기 위해 옵션의 통신 인터페이스(60)가 구비될 수도 있다. 따라서, 인터페이스(60)는 적외선 링크, 모뎀, 네트워크 카드 등의 각종 형태의 통신장치를 구비할 수 있다.

[0039] 모바일폰(30)은 마이크(29), 아날로그/디지털(A/D) 변환기(37), 및 저장부(54)에 저장된 옵션의 스피치 인식 프로그램을 구비한다. 장치(30)의 사용자로부터의 가청 정보, 명령 또는 커맨드에 응답하여, 마이크(29)는 A/D 변환기(37)에 의해 디지털변환된 스피치 신호를 제공한다. 스피치 인식 프로그램은 디지털변환된 스피치 신호에 대해 정규화 및/또는 특징추출 기능을 수행하여 스피치 인식의 중간결과를 획득한다. 스피치 데이터는 무선 트랜시버(52) 또는 통신 인터페이스(60)를 이용하여 도 5의 아키텍처에 도시된 리모트 스피치 서버(204)[이에 대해서는 후술하기로 함]에 송신된다. 그런 다음, 인식 결과는 모바일 장치(30)에서의 렌더링(예컨대, 비주얼 및/또는 가청) 및 웹서버(202)로의 이벤트 송신을 위해 모바일 장치(30)에 리턴되며, 여기서 웹서버(202)와 모바일 장치(30)는 클라이언트/서버 관계로 동작한다.

[0040] 도 3은 2G 폰(81)의 실시예를 평면도로 나타낸 것이다. 폰(81)은 디스플레이(82) 및 키패드(84)를 구비한다. 일반적으로, 폰(81)은 음성 채널(87로 도시됨)을 통한 음성 통화를 행하고 데이터 채널(85로 도시됨)을 통한 디지털 데이터의 송신 및 수신을 행하기 위한 회로를 구비한다. 이러한 타입의 2G 폰은 수많은 제조사로부터 구입할 수 있으며 잘 정의되어 있는 표준 및 프로토콜에 따라서 동작한다. 상기한 회로의 동작과 관련한 상세에 대해서는 본 발명을 이해하는데 필요하지 않을 것으로 생각되므로 생략하기로 한다.

[0041] 본 발명은 전술한 휴대용 또는 모바일 컴퓨팅 장치뿐만 아니라 범용 데스크톱 컴퓨터와 같은 수많은 다른 컴퓨팅 장치에서도 이용될 수 있다. 예를 들어, 문자 숫자 조합이 모두 구비된 키보드와 같은 다른 통상의 입력 장치로는 조작하기 힘든 경우에도, 상기한 아키텍처(200)는 육체적으로 능력이 제한된 사용자도 컴퓨터 또는 기타 컴퓨팅 장치에 텍스트를 입력할 수 있도록 한다.

[0042] 아래에서는 도 4에 도시된 범용 컴퓨터(120)에 대해 간략히 설명하기로 한다. 그러나, 컴퓨터(120)는 적절한 컴퓨팅 환경의 일례일 뿐이며 본 발명의 용도나 기능의 범위가 이에 국한되는 것은 아니다. 또한, 상기한 컴퓨터(120)는 예시된 컴포넌트들 중 임의의 하나 또는 조합에 관하여 종속되거나 필수조건인 것으로 해석되어서는 안된다. 또한, 퍼스널 컴퓨터(120)도 웹서버(202), 스피치 서버(204), 텔레포니 음성 브라우저(212) 등 - 이에 국한되지 않음 - 과 같은 아키텍처(200)의 다른 컴포넌트에 대한 적절한 오퍼레이팅 환경을 제공할 수 있다.

[0043] 본 발명은 컴퓨터에 의해 실행되는 프로그램 모듈과 같은 일반적인 의미의 컴퓨터 실행가능 명령으로 기술될 수 있다. 일반적으로, 프로그램 모듈은 특정 태스크를 수행하거나 특정 추상 데이터타입을 구현하는, 루틴, 프로그램, 오브젝트, 컴포넌트, 데이터 구조 등을 포함한다. 본 발명은 또한 복수의 태스크가 통신 네트워크를 통해 링크되어 있는 리모트 처리장치들에 의해 수행되는, 분산 컴퓨팅 환경에서 실시될 수도 있다. 분산 컴퓨팅 환경에서는 프로그램 모듈이 메모리 저장장치를 구비한 로컬 및 리모트 컴퓨터 저장매체에 위치할 수 있다. 프로그램 및 모듈에 의해 수행되는 태스크에 대해서는 도면을 참조하여 후술하기로 한다. 본 기술분야의 숙련된 자들이라면 상기한 설명 및 도면을 임의의 형태의 컴퓨터 판독가능 매체에 기록될 수 있는 프로세서 실행가능 명령으로서 구현할 수 있을 것이다.

[0044] 도 4를 참조하면, 컴퓨터(120)의 컴포넌트는 이것에 국한되지는 않지만 처리부(140), 시스템 메모리(150) 및 시스템 버스(141)를 포함하며, 상기한 시스템 버스(141)는 시스템 메모리를 포함한 각종 시스템 컴포넌트를 처리부(140)에 연결시킨다. 시스템 버스(141)는 각종 버스 구조 중 임의의 것을 이용한 메모리 버스 또는 메모리 컨트롤러, 주변 버스, 및 로컬 버스를 포함한 수개의 버스 구조 중 임의의 것일 수 있다. 이것에 국한되지는 않지만 예를 들어, 이러한 구조에는 ISA(Industry Standard Architecture) 버스, USB(Universal Serial Bus), MCA(Micro Channel Architecture) 버스, EISA(Enhanced ISA) 버스, VESA(Video Electronics Standards Association) 로컬 버스 및 PCI(Peripheral Component Interconnect) 버스 - Mezzanine 버스로도 알려져 있음 - 를 포함한다. 컴퓨터(120)는 각종의 컴퓨터 판독가능 매체를 포함하는 것이 일반적이다. 컴퓨터 판독가능

매체는 컴퓨터(120)에 의한 액세스가 가능한 임의의 매체일 수 있으며, 활성 및 불활성 매체, 착탈식 및 비착탈식 매체를 포함한다. 이것에 국한되지는 않지만 예를 들어, 컴퓨터 판독가능 매체는 컴퓨터 저장매체 및 통신 매체를 구비할 수 있다. 컴퓨터 저장매체는 컴퓨터 판독가능 명령, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보를 저장하기 위한 임의의 방법 또는 기술에 의해 구현된 활성 및 불활성, 착탈식 및 비착탈식 매체를 포함한다. 컴퓨터 저장매체로는 이것에 국한되지는 않지만 RAM, ROM, EEPROM, 플래시 메모리 또는 기타 메모리 테크놀로지, CD-ROM, DVD(digital versatile disk) 또는 기타 광디스크 스토리지, 마그네틱 카세트, 마그네틱 테이프, 마그네틱 디스크 스토리지 또는 기타 마그네틱 스토리지 장치, 또는 원하는 정보를 저장하고 또한 컴퓨터(120)에 의한 액세스가 가능한 임의의 다른 매체를 포함할 수 있다.

[0045] 통신매체는 일반적으로 컴퓨터 판독가능 명령, 데이터 구조, 프로그램 모듈, 또는 반송파나 기타 전송 매체 커리어 등의 변조 데이터 신호인 기타 데이터를 구현하고 있으며 임의의 정보전달매체를 포함한다. 여기서, "변조 데이터 신호"는 하나 이상의 특징세트를 갖는 신호 또는 그 신호 내의 정보를 인코딩하는 방식으로 변경된 신호를 의미한다. 이것에 국한되지는 않지만 예를 들어, 통신매체는 유선 네트워크 또는 직접접속 네트워크 등의 유선매체와 어쿠스틱, FR, 적외 및 기타 무선매체 등의 무선매체를 포함한다. 상기한 것들의 임의의 조합도 컴퓨터 판독가능 매체의 범위에 포함될 수 있다.

[0046] 시스템 메모리(150)는 ROM(151) 및 RAM(152)와 같은 활성 및/또는 불활성 메모리 형태의 컴퓨터 저장매체를 포함한다. 상기한 ROM(151)에는 일반적으로 기동시 등에 컴퓨터(120) 내의 각 엘리먼트들 사이의 정보 교환을 도와주는 BIOS(basic input/output system)(153)가 저장되어 있다. RAM(152)은 일반적으로 처리부(140)에 의해 현재 동작되고 있는 및/또는 직접 액세스가 가능한 데이터 및/또는 프로그램 모듈을 저장하고 있다. 이것에 국한되지는 않지만 예를 들어, 도 4는 오퍼레이팅 시스템(54), 애플리케이션 프로그램(155), 기타 프로그램 모듈(156) 및 프로그램 데이터(157)를 도시하고 있다.

[0047] 컴퓨터(120)는 기타 착탈식/비착탈식 활성/불활성 컴퓨터 저장매체를 포함할 수도 있다. 이것에 국한되지는 않지만 예를 들어, 도 4는 비착탈식 불활성 마그네틱 매체에의 판독 또는 기록을 수행하는 하드디스크 드라이브(161), 착탈식 불활성 마그네틱 디스크(172)에의 판독 또는 기록을 수행하는 마그네틱 디스크 드라이브(171), 및 착탈식 불활성 광디스크(176)(예컨대 CD-ROM, 또는 기타 광학 매체)에의 판독 또는 기록을 수행하는 광디스크 드라이브(175)를 도시하고 있다. 예시적 오퍼레이팅 환경에서 사용될 수 있는 다른 착탈식/비착탈식 활성/불활성 컴퓨터 저장매체로는 이것에 국한되지는 않지만 마그네틱 테이프 카세트, 플래시 메모리 카드, DVD, 디지털 비디오 테이프, 솔리드 스테이트 RAM, 솔리드 스테이트 ROM 등을 포함한다. 하드디스크 드라이브(161)는 일반적으로 인터페이스(160)와 같은 불활성 메모리 인터페이스를 통해 시스템 버스(141)에 접속되며, 마그네틱 디스크 드라이브(171) 및 광디스크 드라이브(175)는 일반적으로 인터페이스(170)와 같은 착탈식 메모리 인터페이스에 의해 시스템 버스(141)에 접속된다.

[0048] 도 4에 도시된 상기한 드라이브 및 그 관련 컴퓨터 저장매체는 컴퓨터 판독가능 명령, 데이터 구조, 프로그램 모듈 및 컴퓨터(120)용의 기타 데이터에 대한 저장부를 제공한다. 도 4에서, 예를 들어 하드디스크 드라이브(161)는 오퍼레이팅 시스템(164), 애플리케이션 프로그램(165), 기타 프로그램 모듈(166) 및 프로그램 데이터(167)를 저장하고 있는 것으로 도시되어 있다. 여기서, 이들 컴포넌트들은 오퍼레이팅 시스템(154), 애플리케이션 프로그램(155), 기타 프로그램 모듈(156) 및 프로그램 데이터(157)와 동일할 수도 있고 상이한 것일 수도 있다. 오퍼레이팅 시스템(164), 애플리케이션 프로그램(165), 기타 프로그램 모듈(166) 및 프로그램 데이터(167)에 대해 다른 번호를 부여한 것은 이들이 상이한 카피일 수도 있음을 나타내기 위함이다.

[0049] 사용자는 키보드(182), 마이크(183) 및 지시장치(181)(예컨대, 마우스, 트랙볼 또는 터치패드 등)와 같은 입력 장치를 통해 컴퓨터(120)에 커맨드 및 정보를 입력할 수 있다. 다른 입력장치(도시되지 않음)로는 조이스틱, 게임패드, 위성수신기, 스캐너 등을 포함한다. 이들 및 다른 입력장치는 일반적으로 시스템 버스에 연결된 사용자 입력 인터페이스(180)를 통해 처리부(140)에 연결되지만, 병렬포트, 게임포트 또는 USB와 같은 다른 인터페이스 및 버스구조에 연결될 수도 있다. 시스템 버스(141)에는 또한 비디오 인터페이스(185)와 같은 인터페이스를 통해 모니터(184) 또는 다른 타입의 표시장치가 연결되어 있다. 모니터 외에도, 컴퓨터는 스피커(187) 및 프린터(186)와 같은 다른 주변출력장치를 구비할 수 있으며, 상기한 주변출력장치는 출력 주변 인터페이스(188)를 통해 접속된다.

[0050] 컴퓨터(120)는 리모트 컴퓨터(194)와 같은 하나 이상의 리모트 컴퓨터와의 논리접속을 이용하는 네트워크 환경에서 동작할 수 있다. 리모트 컴퓨터(194)는 퍼스널 컴퓨터, 핸드헬드 장치, 서버, 라우터, 네트워크 PC, 피어 장치 또는 커먼(common) 네트워크 노드일 수 있으며, 일반적으로 컴퓨터(120)와 관련하여 기술한 바 있는 다수



또는 모든 엘리먼트를 포함한다. 도 4에 도시된 논리 접속부는 LAN(191) 및 WAN(193)을 포함하며 다른 네트워크를 포함할 수도 있다. 이러한 네트워크 환경은 사무실, 거대 컴퓨터 네트워크, 인트라넷 및 인터넷에서 흔히 볼 수 있다.

[0051] LAN 네트워크 환경에서 사용하는 경우, 컴퓨터(120)는 네트워크 인터페이스 또는 어댑터(190)를 통해 LAN(191)에 연결된다. WAN 네트워크 환경에서 사용하는 경우, 컴퓨터(120)는 일반적으로 인터넷과 같은 WAN(193)을 통한 통신을 확립하기 위한 모뎀 또는 기타 수단을 포함한다. 내장형 또는 외장형의 모뎀(192)은 사용자 입력 인터페이스(180) 또는 기타 적절한 메커니즘을 통해 시스템 버스(141)에 연결된다. 네트워크 환경에서, 컴퓨터(120)와 관련하여 도시한 컴퓨터 모듈 또는 그 일부는 리모트 메모리 저장장치에 저장될 수 있다. 이것에 국한되지는 않지만 예를 들어, 도 4에서는 리모트 애플리케이션 프로그램(195)이 리모트 컴퓨터(194)에 위치하는 것으로 도시하고 있다. 여기서, 도시된 네트워크 접속부는 예시를 위한 것이며 컴퓨터들 사이에 통신을 확립하기 위한 다른 수단이 사용될 수도 있다.

[0052] 도 5는 본 발명에 이용될 수 있는 웹기반 스피치 인식을 위한 아키텍처(200)를 도시한 것이다. 전술한 바와 같이, 웹서버(202)에 저장된 정보는 모바일 장치(30), 심플 폰(80) 또는 2G 폰(81)을 통해 액세스될 수 있다. 상기한 아키텍처(200) 및 이것에 사용된 마크업 랭기지에 대해서는 미국특허출원 제2002-0169806호(2002.11.14)에 보다 상세히 설명되어 있다.

[0053] 일반적으로, 장치(30)부터 살펴보면, 장치(30)는 웹서버(202)에 의해 제공된 HTML+ 스크립트 등을 실행한다. 음성 인식이 필요한 경우 전술한 바와 같이 장치(30)에 의해 전처리된 디지털변환된 오디오 신호일 수 있는 스피치 데이터 또는 스피치 특징이 스피치 인식 중에 사용할 문법 또는 랭기지 모델의 표시와 함께 스피치 서버(204)에 제공된다. 스피치 서버(204)의 구현은 다양한 형태를 취할 수 있지만 그 중의 하나만을 도시하고 있으며, 일반적으로 음성 인식기(211)를 포함한다. 스피치 인식의 결과는 필요한 경우 로컬 렌더링을 위해 상기한 장치(30)에 리턴되어 제공된다. 임의의 GUI(graphic user interface) - 사용된 경우 - 및 음성 인식을 통한 정보의 컴파일(compilation) 과정 중에, 장치(30)는 필요한 경우 추가의 처리 및 추가의 HTML 스크립트 수신을 위해 상기한 정보를 웹서버(202)에 송신한다.

[0054] 도 5에 도시된 바와 같이, 장치(30), 2G 폰(81), 웹서버(202), 텔레포니 음성 브라우저(212) 및 스피치 서버(204)는 공통 연결되어 있으며 네트워크(205)를 통해 개별적으로 어드레스할 수 있다(상기한 네트워크는 도 5에서 인터넷과 같은 WAN임). 따라서 이들 장치가 서로 물리적으로 인접하여 배치될 필요는 없다. 특히 웹서버(202)가 스피치 서버(204)를 구비하고 있을 필요는 없다. 이와 같이, 웹서버(202)에서의 프로그래밍(authoring)에 있어서는, 프로그래머(author)가 스피치 서버(204)의 상세(intricacies)를 알고 있을 필요가 없으며 자신이 의도하는 애플리케이션에 프로그래밍(authoring)을 집중할 수 있다. 오히려, 스피치 서버(204)는 독립 설계가 가능하며 네트워크(205)에 독립적으로 연결될 수 있으며, 따라서 웹서버(202)에서의 추가 변경의 필요없이 업데이트 및 개량이 가능하게 된다. 또한 스피치 서버(204)는 다수의 클라이언트 장치(30), 폰(80 및 81) 및/또는 웹서버(202)를 서비스할 수 있다.

[0055] 또다른 실시예에 있어서, 웹서버(202), 스피치 서버(204) 및 클라이언트(30)는 구현 머신의 능력에 따라서 결합 구성될 수 있다. 예를 들어, 클라이언트가 범용 컴퓨터, 예컨대 퍼스널 컴퓨터를 구비하고 있다면 클라이언트는 스피치 서버(204)를 포함하고 있을 것이다. 마찬가지로, 필요하다면 웹서버(202)와 스피치 서버(204)를 하나의 머신에 통합할 수도 있다.

[0056] 클라이언트 장치(30)에 관하여, 클라이언트/서버 시스템에서의 음성 인식 처리 방법은, 서버(202)로부터 클라이언트 장치의 사용자로부터 스피치 데이터를 획득하도록 설정된 확장자를 갖는 마크업 랭기지 페이지를 수신하는 단계; 마크업 랭기지 페이지를 클라이언트 장치에서 실행하는 단계; 클라이언트로부터 원격지에 있는 스피치 서버에 [사용자로부터 획득한 스피치를 나타내는] 스피치 데이터 및 관련 문법을 송신하는 단계; 및 스피치 서버로부터의 인식 결과를 클라이언트가 수신하는 단계를 포함한다. 클라이언트/서버 시스템 내의 클라이언트 장치에서의 실행을 위한 마크업 랭기지를 갖는 컴퓨터 판독가능 매체가 제공될 수 있으며, 상기한 마크업 랭기지는 클라이언트 장치를 통해 입력된 스피치와 연관될 문법을 지시하는 명령을 포함한다.

[0057] 폰(80)을 통한 웹서버(202)에의 액세스는, 폰(80)을 유선 또는 무선 전화 네트워크(208)에 연결하고 또한 폰(80)을 3rd 파티 게이트웨이(210)에 연결하는 것을 포함한다. 게이트웨이(210)는 폰(80)을 텔레포니 음성 브라우저(212)에 연결시킨다. 텔레포니 음성 브라우저(212)는 텔레포니 인터페이스 및 음성 브라우저(216)를 제공하는 미디어 서버(214)를 포함한다. 장치(30)와 마찬가지로, 텔레포니 음성 브라우저(212)는 웹서버(202)로부터 HTML 스크립트 등을 수신한다. 여기서 중요한 것은, HTML 스크립트가 장치(30)에 제공된 것과 유사한 형태

로 되어 있다는 것이다. 이와 같이, 웹서버(202)는 장치(30) 및 폰(80)을 개별적으로 서포트하거나 표준 GUI 클라이언트를 개별적으로 서포트할 필요가 없다. 오히려, 커먼(common) 마크업 랭기지가 사용될 수 있다. 또한, 장치(30)와 마찬가지로, 폰(80)에 의해 송신된 가정 신호에 대한 음성 인식은 네트워크(205)를 통하거나 또는 예컨대 TCP/IP를 이용한 전용선(207)을 통해 음성 브라우저(216)로부터 스피치 서버(204)에 제공된다. 인식 결과 및 기타 정보는 텔레포니 음성 브라우저(212) 및 폰(80)을 통해 사용자에게 청취가능하도록 렌더링되어 리턴된다.

[0058] 상기한 바와 같이, HTML, XHTML, cHTML, XML, WML, 또는 임의의 다른 SGML 파생 마크업과 같은 마크업 랭기지는 클라이언트/서버 아키텍처에서 스피치 인식을 제공하는 컨트롤 및/또는 오브젝트를 포함할 수 있다. 이와 같이, 프로그래머는 이러한 아키텍처에서 이용되는 지배적인 웹개발 플랫폼인 이들 마크업 랭기지에서의 모든 툴과 전문지식을 이용할 수 있게 된다.

[0059] 일반적으로, 컨트롤 및/또는 오브젝트에는 하기의 평선 중 하나 이상을 포함한다: 인식기 설정, 실행 및/또는 후처리를 위한 인식기 컨트롤 및/또는 오브젝트; 합성기 설정 및 프롬프트 플레이를 위한 합성기 컨트롤 및/또는 오브젝트; 입력 문법 리소스 규정을 위한 문법 컨트롤 및/또는 오브젝트; 및 인식결과의 처리를 위한 바인딩 컨트롤 및/또는 오브젝트. 이러한 확장은 경량 마크업 레이어가 되도록 하기 위해 고안된 것으로, 기존의 마크업 랭기지에 스피치 인터페이스의 능력을 추가한 것이다. 따라서, 이러한 확장은 자신이 속한 하이레벨 페이지(예컨대 HTML); 상기 확장이 언어관련(linguistic) 리소스를 참조하는데 이용하는 로우레벨 포맷(예컨대 텍스트 투 스피치 및 문법 포맷); 및 스피치 서버(204)에서 이용하는 스피치 합성 플랫폼 및 상기한 인식의 개별 속성에 독립적으로 존재한다.

[0060] 여기서 주목할 점은, 본 발명은 SALT(speech application language tags)와 같은 마크업 랭기지 확장자를 이용하여 구현될 수 있다는 것이다. SALT는 예를 들어 퍼스널 컴퓨터, 전화, 테블릿 PC 및 무선 모바일 장치로부터 정보, 애플리케이션 및 웹서버로의 액세스를 가능하게 하기 위한 개발 표준이다. SALT는 HTML, XHTML 및 XML 등의 기존 마크업 랭기지에도 확장된다. SALT 1.0 사양은 온라인(<http://www.SALTforum.org>)으로 찾을 수 있다.

[0061] 전술한 바와 같이 아키텍처(200)에는 2G 폰(81)을 통한 멀티모드 인터랙션이 제공된다. 일반적으로, 멀티모드 인터랙션은 사용자의 희망에 따라 자연스런 방식으로 웹서버(202)로부터 정보의 액세스를 허용한다. 특히, 키패드 조작에 의한 텍스트 형태의 커멘트를 제공하고 그 결과를 비주얼 표시 텍스트로서 수신하는 것이 아니라, 사용자는 입력 매체로서 스피치를 제공하여 그 결과를 비주얼 또는 원하는 경우 합성된 스피치로서 수신할 것을 선택할 수 있다. 그러나, 2G 폰(81)과 같은 장치는 제한된 처리 능력과 잘 알려진 바와 같이 요구조건을 갖기 때문에, 인터넷과 같은 네트워크에의 연결을 위한 데이터 채널이 존재하고 통화를 행하기 위한 별도의 음성 채널을 구비되더라도, 이들 채널은 동시에 액세스될 수 없다. 그 결과, 데이터 및 음성 채널을 필요로 하는 멀티모드 인터랙션이 순차 멀티모드로서 알려진 바와 같이 순차적으로 수행되어야 한다. 그러나, 전술한 아키텍처(200) 및 후술하게 될 방법은 웹서버(202)와의 순차 멀티모드 인터랙션을 제공하는데 이용할 수 있다. 상기한 아키텍처에 2G 폰(81)을 통합하는 경우, 웹서버(202)에의 액세스가 장치(30) 또는 폰(80)과 같은 다른 장치와 일치(consistent)하므로, 장치(30) 및 폰(80)에 추가하여 2G 폰(81)을 서포트하기 위하여 웹서버(202) 및 여기서 동작하는 애플리케이션을 근본적으로(drastically) 변경할 필요없이 특히 유용하다. 이와 같이, 애플리케이션 개발자는 정보를 액세스할 수 있는 각각의 장치를 서포트하기 위하여 개별 애플리케이션을 제공해야 하는 부담을 덜 수 있을뿐만 아니라, 능력이 서로 다른 다수의 상이한 장치들을 서포트할 수 있는 보다 일원화된 코드를 제공할 수 있게 된다.

[0062] 도 6은 2G 폰(81)에 적용될 수 있는 순차 멀티모드 시나리오를 나타낸 것으로, 스피치 인식 결과가 WML/XHTML 페이지를 이용하여 텍스트 형태로 WAP를 통해 제시된다.

[0063] 무선 애플리케이션 프로토콜(WAP)은 공개된 공지의 사용으로서, 이는 사용자로 하여금 모바일폰을 통해 정보를 액세스하고 모바일폰의 디스플레이(82)에 콘텐츠 및 심플 그래픽을 표시할 수 있도록 한다. WAP는 음성에 의한 인터랙트 능력이 결여되어 있으며 또한 그 입력은 일반적으로 대부분의 모바일폰에서 12개의 키로 제한된다.

[0064] 알려진 바와 같이, 2G 폰(81)에서도 전세계에서 채용되고 있는 모바일 서비스인 SMS(단문 메시지 서비스)를 지원하며 이 서비스에 의해 무선장치에 문자 숫자 조합의 메시지를 송신할 수 있다.

[0065] 도 7a 및 도 7b는 2G 폰(81)으로 순차 멀티모드 스피치 인식을 수행하기 위한 방법(300)의 예시적 단계들을 나타낸 것이다.

- [0066] 도시된 예에서, 단계 304에서는 항공 예약을 위한 애플리케이션에 액세스하기 위해 도 6의 화살표 302로 표시된 최초 요청이 웹서버(202)에 전달되는 것으로 가정한다.
- [0067] 단계 306에서 웹서버(202)는 2G 폰(81)에 소정의 페이지를 제공하며(화살표 307), 본 예에서는 상기 페이지는 출발 도시의 지정을 위한 텍스트박스 또는 데이터 필드 입력을 위한 다른 표시뿐만 아니라, 출발 주(state)의 지정을 위한 텍스트박스 또는 데이터 필드 입력을 위한 다른 표시를 포함한다. 이들 필드는 도 8에서 308 및 310으로 도시되어 있다. 상기한 웹페이지는 무선 WAP/SMS 데이터 채널(85)을 통해 웹서버로부터 2G 폰으로의 송신된다.
- [0068] 종래의 2G 폰에서는 사용자가 상기한 각 텍스트박스 또는 데이터 필드(308, 310)에 텍스트를 입력하기 위한 옵션을 구비하고 있었다. 그러나, 2G 폰에서는 사용자가 이용할 수 있는 키패드(84) 상의 키가 12개로 제한되는 것이 일반적이므로, 각각의 커먼 문자 숫자 조합 심볼을 제공하기 위해 조작해야 한다.
- [0069] 본 발명에서는 사용자가 각 데이터 필드(308, 310)에 따른 스피치 입력을 제공할 수 있으므로 제한된 키패드(84)를 조작하는 번거로움을 배제할 수 있다.
- [0070] 단계 312에서 사용자는 스피치 입력을 제공할 것이라는 표시를 제공한다. 이 표시는 키패드(84) 중 하나의 키를 누르거나 2G 폰(81) 상의 특정 버튼(89)을 누르는 형태로 제공될 수 있다. 그러나, 다른 형태의 표시로서 2G 폰(81)내에서 처리 및 인식이 가능한 선정된 음성 커맨드를 포함할 수 있다.
- [0071] 단계 314에서 2G 폰(81)은 도 6의 화살표 316으로 표시된 바와 같이 텔레포니 음성 브라우저(212)와 음성 통화를 시발한다. 단계 318에서 텔레포니 음성 브라우저(212)에 접속된 다음, 텔레포니 음성 브라우저(212)는 단계 306에서 이전에 송신된 웹페이지에 따라서 상기한 웹서버(202)로부터 음성 인식을 위한 태그 결합된 스피치-인에이블 웹페이지를 요청한다. 이는 화살표 320으로 표시되어 있다. 일 실시예에서, 화살표 323으로 표시된 바와 같이, 단계 321에서 텔레포니 음성 브라우저(212)에 제공되게 될 올바른 웹페이지는 웹서버(202)에 의해 2G 폰(81) 상의 상기한 페이지와 결합된 폰번호 또는 다른 심볼 식별자를 통해 확정되게 된다. 그러면, 웹서버(202)는 상기한 폰번호 또는 기타 식별자를 통해 데이터 채널(85)을 통해 2G 폰(81)에 직접 송신된 올바른 페이지들의 결합뿐만 아니라, 웹서버(202)와 텔레포니 음성 브라우저(212) 사이에 송신된 페이지들을 보유할 수 있게 된다. 웹서버(202)로부터 텔레포니 음성 브라우저(212)에 송신된 페이지는 단계 306의 웹페이지에 대해 2G 폰(81)에 송신된 데이터 필드의 스피치 인식에 필요한 모든 문법 또는 표시를 구비하게 된다.
- [0072] 상기한 텔레포니 음성 브라우저(212)가 음성 채널(87)을 통해 사용자로부터의 스피치를 수신할 수 있는 경우에는, 단계 324에서 사용자는 상기한 필드에 대한 스피치를 제공한다. 여기서 주목할 점은, 일 실시예에서는 사용자의 스피킹 개시를 프롬프트하기 위하여 톤 또는 음성 커맨드와 같은 적절한 프롬프트가 텔레포니 음성 브라우저(212)에 의해 사용자에게 제공될 수 있다는 것이다. 텔레포니 음성 브라우저(212)는 웹서버(202)로부터 대응하는 스피치-인에이블 페이지를 수신한 대에 상기한 프롬프트를 시발할 수 있다. 그러나 다른 실시예에서는 텔레포니 음성 브라우저(212)가 상기한 스피치-인에이블 웹페이지의 수신 전에 상기한 프롬프트를 제공하고, 단계 312에서 사용자가 스피치의 제공을 표시한 때부터 단계 324에서 스피치를 실제로 제공할 때까지의 시간을 최소화하기 위하여, 수신한 스피치를 적절한 버퍼 또는 기타 저장 장치에 일시 저장하고 있다.
- [0073] 입력된 스피치는 스피치 서버(202)를 이용하여 폰(80)에 의한 동작과 관련하여 전송한 것과 대체적으로 동일한 방식으로 처리된다. 특히 텔레포니 음성 브라우저(212)는 화살표 328로 표시된 바와 같이 단계 326에서 입력된 스피치를 스피치 서버(204)에 제공한다. 스피치 서버(204)는 인식을 수행하고, 그 결과는 도 6의 화살표 332로 표시된 바와 같이 단계 330에서 웹서버(202)에 재송신된다.
- [0074] 웹서버(202)는 2G 폰(81)과 연관된 폰번호 또는 기타 식별자에 따라서 인식된 결과를 수신한다. 단계 340에서 웹서버(202)는 도 6의 화살표 342로 표시된 바와 같이 데이터 채널(85)을 이용하여 SMS 메시지를 2G 폰(81)에 제공한다. 상기한 SMS 메시지가 인식 결과를 포함하고 있을 수도 있지만, 다른 실시예에서는 상기한 SMS 메시지가 웹페이지의 어드레스(예컨대 URL 링크)로 구성된다. SMS 메시지를 수신하게 되면, 단계 344에서 사용자 또는 사용자 에이전트는 데이터 채널을 분명하게 또는 불분명하게 이용하여, 화살표 346으로 표시된 바와 같이, 인식 결과를 포함하는 업데이트된 웹페이지를 검색하며, 텔레포니 음성 브라우저(212)에의 음성 통화는 차단된다. 또다른 실시예에서, 상기한 음성 채널의 차단은 텔레포니 서버가 음성 브라우저 페이지 상의 모든 필요한 정보를 수집한 후에 상기 텔레포니 서버에 의해 수행될 수 있다.
- [0075] 다음으로, 단계 348에서 웹서버(202)는 화살표 350으로 표시된 바와 같이 인식 결과를 구비한 2G 폰(81)에 새로운 페이지를 제공한다. 도 9는 사용자가 제공한 스피치 입력에 기초하여 스피치 인식 결과가 추가된 텍스트박

스(308 및 310)를 나타낸 것이다.

[0076] 이상 2G 폰에서의 스피치 입력을 제공하는 순차 멀티모드 동작을 설명하였다. 도 5에 도시된 아키텍처는 도 7a 및 도 7b에 도시된 방법의 각 동작을 반복 수행함으로써, 2G 폰(81)의 제한된 능력 하에서 효과적인 스피치 인터랙션을 제공하기 위하여 웹페이지에 연관된 다른 필드, 또는 다른 웹페이지에 연관된 필드들에 대한 스피치 입력을 제공할 수 있도록 한다.

[0077] 이상, 본 발명을 특정 실시예에 관하여 설명하였지만, 본 기술분야의 숙련된 자들이라면 본 발명의 사상 및 범주를 이탈하지 않는 범위 내에서 그 형태나 상세에 있어서 각종의 변경이 가능함을 이해할 것이다.

### 발명의 효과

[0078] 전술한 바와 같이, 본 발명에 따르면, 서버/클라이언트 아키텍처에서 특히 2G 폰과 같은 장치에 대한 서버 정보를 액세스하는데 유용한 아키텍처 및 방법이 제공된다.

### 도면의 간단한 설명

[0001] 도 1은 컴퓨팅 장치의 오퍼레이팅 환경을 나타낸 평면도.

[0002] 도 2는 도 1의 컴퓨팅 장치의 블록도.

[0003] 도 3은 2G 휴대폰의 평면도.

[0004] 도 4는 범용 컴퓨터의 블록도.

[0005] 도 5는 클라이언트/서버 시스템의 아키텍처를 나타낸 블록도.

[0006] 도 6은 순차 멀티모드 인터랙션을 제공하도록 도 5의 아키텍처의 구성요소들에 제공된 접속을 나타낸 블록도.

[0007] 도 7a 및 도 7b는 순차 멀티모드 인터랙션을 제공하기 위한 방법의 일례를 설명하기 위한 흐름도.

[0008] 도 8은 2G 폰에 렌더링되는 텍스트 박스의 일례를 나타낸 도면.

[0009] 도 9는 2G 폰에 인식 결과가 렌더링된 텍스트 박스의 일례를 나타낸 도면.

[0010] 〈도면의 주요부분에 대한 부호의 설명〉

[0011] 30: 클라이언트

[0012] 80: 심플 폰

[0013] 81: 2G 폰

[0014] 202: 웹서버

[0015] 204: 스피치 서버

[0016] 205: PSDN

[0017] 206: 웹페이지

[0018] 210: 3rd 파티 VOIP 게이트웨이

[0019] 212: 텔레포니 음성 브라우저

[0020] 214: 미디어 서버

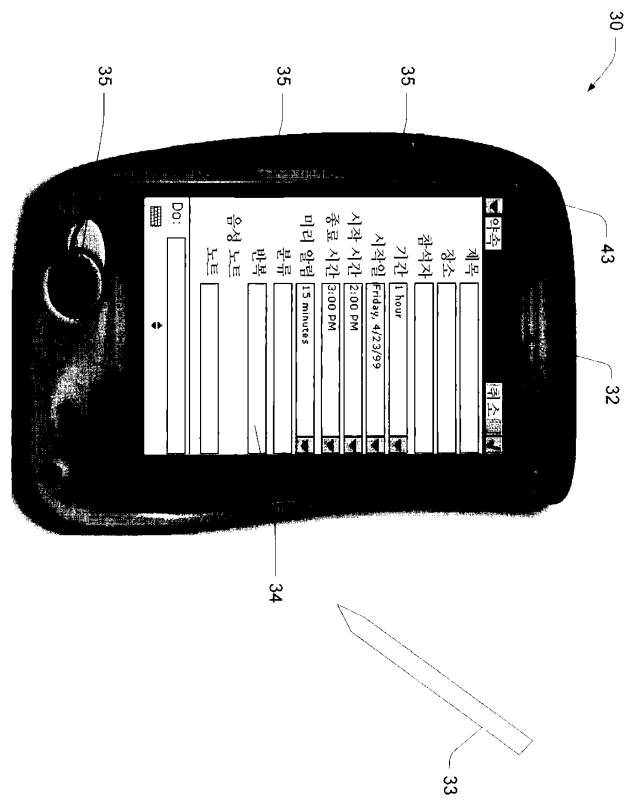
[0021] 216: 음성 브라우저

[0022] 220: 랭기지 모델

[0023] 220: 랭기지 모델

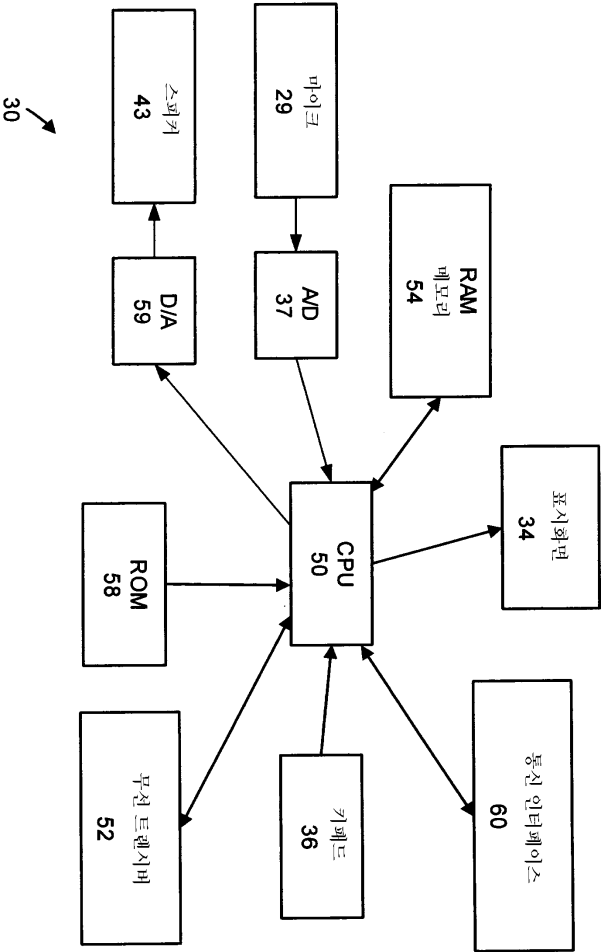
도면

도면1

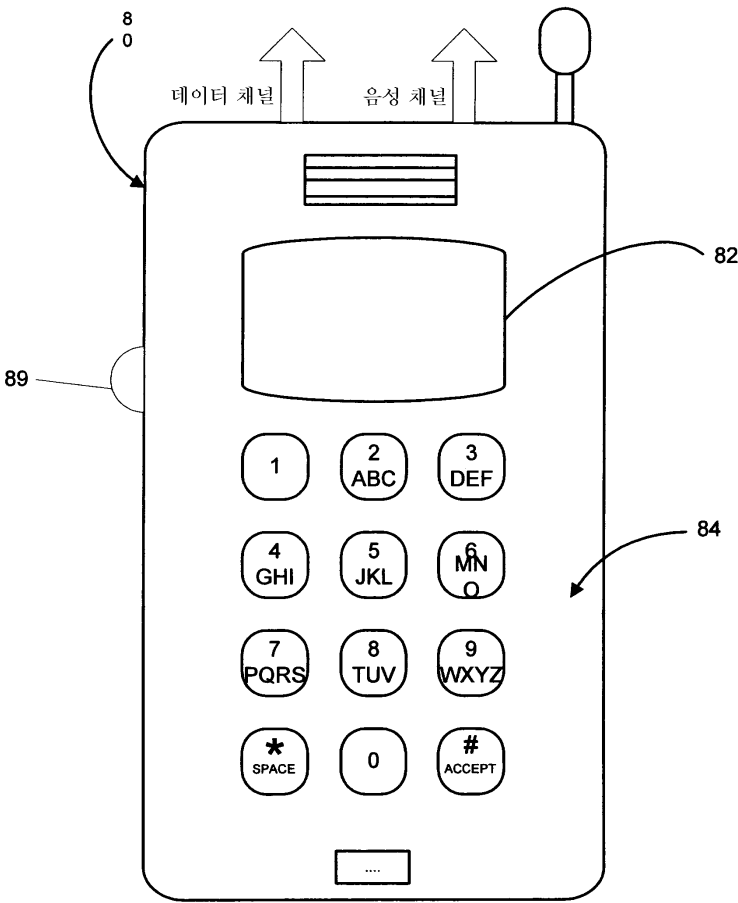




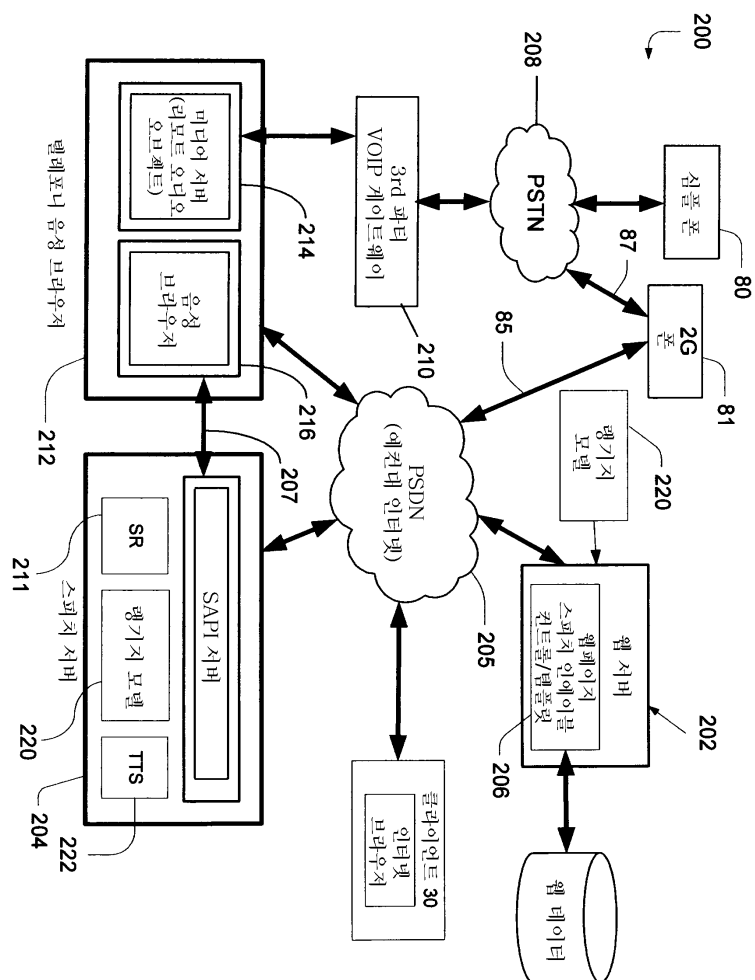
도면2



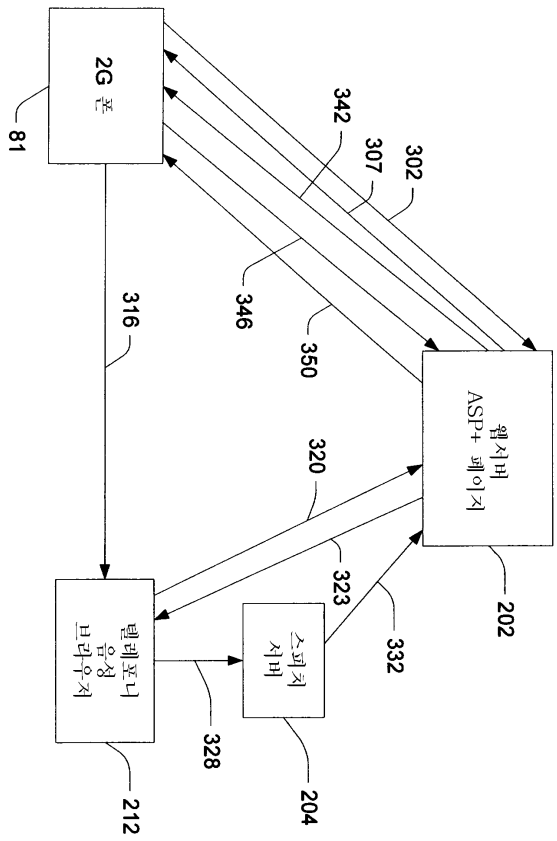
도면3





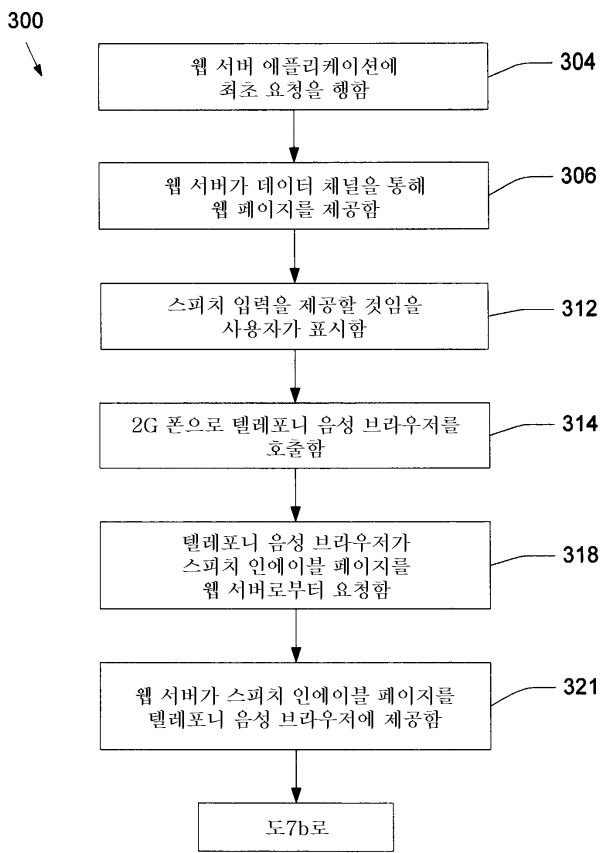


도면6

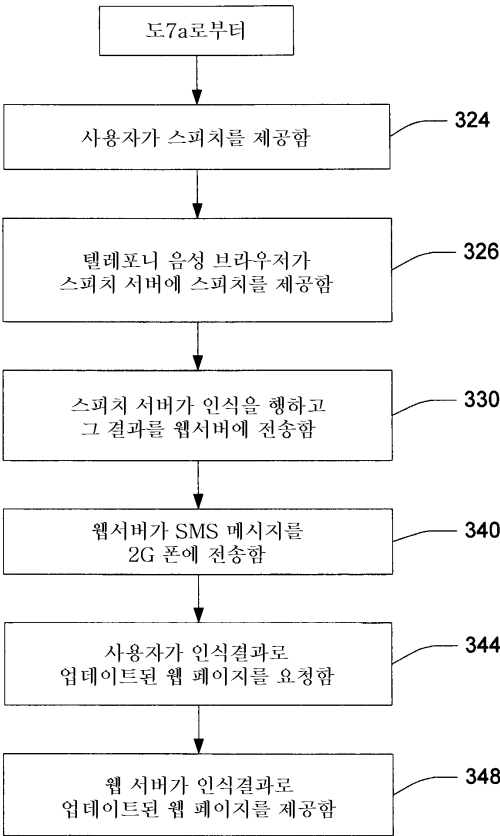




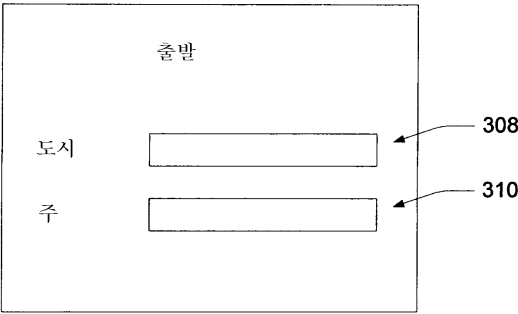
도면7a



도면7b



도면8



도면9

