



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2016년05월16일
(11) 등록번호 10-1620953
(24) 등록일자 2016년05월09일

(51) 국제특허분류(Int. Cl.)

G06F 21/36 (2013.01)

(21) 출원번호 10-2014-0166261

(22) 출원일자 2014년11월26일

심사청구일자 2014년11월26일

(56) 선행기술조사문헌

KR101363668 B1*

US20130145441 A1

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

(72) 발명자

차성덕

서울특별시 성북구 동소문로34길 24, 106동 710호(돈암동, 삼성아파트)

권신일

서울 동대문구 무학로45길 37, 309호 (용두동, 신성빌딩)

(74) 대리인

김동용, 김홍석

전체 청구항 수 : 총 13 항

심사관 : 서광훈

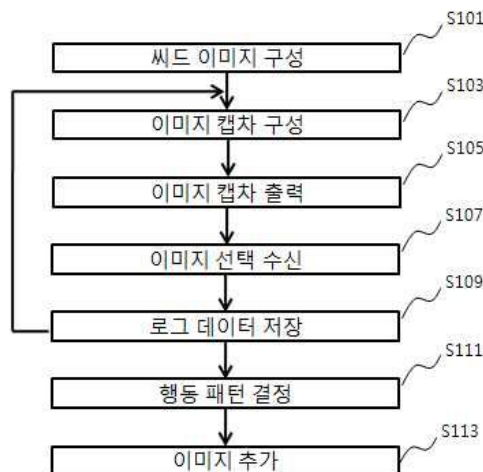
(54) 발명의 명칭 이미지 캡차 제공 방법 및 이미지 캡차 제공 서버

(57) 요약

본 발명은 복수의 이미지로 구성된 이미지 캡차를 출력하는 단계 및 출력된 이미지 캡차의 이미지 선택을 수신하는 단계를 포함하며 복수의 이미지는 접속대상의 이미지 캡차 문제의 통과 결정에 이용되는 제1 그룹의 이미지 및 접속대상의 이미지 캡차 문제의 통과 결정에 이용되지 않는 제2 그룹의 이미지를 포함하는 이미지 캡차 제공 방법 및 이미지 캡차 제공 서버에 관한 것이다.

본 발명을 이용함으로써 이미지의 신분 노출을 방지하고 이미지의 의미 해석에 기초한 이미지 캡차로 접속대상이 사람인지 판별하고 신규의 이미지를 이미지 캡차에 자동으로 포함시킬 수 있도록 하는 효과가 있다.

대표도 - 도3



이 발명을 지원한 국가연구개발사업

과제고유번호 1711015820

부처명 미래창조과학부

연구관리전문기관 정보통신산업진흥원

연구사업명 대학 IT연구센터 육성 지원사업

연구과제명 고품질 융합 소프트웨어 개발지원 도구 연구

기 여 율 1/1

주관기관 고려대학교 산학협력단

연구기간 2014.01.01 ~ 2014.12.31

명세서

청구범위

청구항 1

(a) 제1 그룹의 이미지와 제2 그룹의 이미지를 포함하는 복수의 이미지로 구성된 이미지 캡차를 출력하는 단계;
 (b) 출력된 상기 이미지 캡차의 이미지 선택을 수신하는 단계;
 (c) 상기 이미지 선택을 나타내는 로그 데이터를 저장하는 단계; 및
 (d) 저장된 복수의 로그 데이터로 행동 패턴을 결정하는 단계를 포함하며,
 상기 제1 그룹의 이미지는 접속대상의 이미지 캡차 문제의 통과 결정에 이용되고,
 상기 행동 패턴은 상기 제1 그룹의 이미지 및 상기 제2 그룹의 이미지 중 하나 이상의 이미지에 대한 선택을 나타내는,
 상기 행동 패턴은 루트 노드로부터 하나 이상의 자식 노드를 가지는 트리로 표현되는,
 이미지 캡차 제공 방법.

청구항 2

삭제

청구항 3

제1항에 있어서,
 상기 루트 노드로부터 지정된 리프 노드까지의 패스가 사람을 나타내는 행동 패턴으로 지정되며, 상기 루트 노드 및 상기 루트 노드의 자식 노드는 이미지에 맵핑되는,
 이미지 캡차 제공 방법.

청구항 4

제3항에 있어서,
 상기 트리의 루트 노드에 맵핑되는 이미지는 상기 자식 노드에 맵핑되는 이미지보다 엔트로피가 높고 상기 엔트로피는 접속대상의 모든 IP 개수와 각 이미지가 출력된 IP 개수를 이용하여 계산되는,
 이미지 캡차 제공 방법.

청구항 5

제3항에 있어서,
 (e) 상기 로그 데이터에 기초하여 노드에 맵핑될 이미지를 추가하는 단계;를 더 포함하며,
 추가되는 이미지는 상기 노드에 맵핑된 이미지와 동일한 분류 기능을 수행할 수 있는,
 이미지 캡차 제공 방법.

청구항 6

제1항에 있어서,
 상기 단계 (c)는 상기 단계 (d)의 수행 이후에 반복적으로 수행되고,
 상기 이미지 캡차 제공 방법은,
 (f) 상기 단계 (d) 이후의 상기 단계 (c)에 의해 저장된 복수의 로그 데이터로 트리로 표현되는 행동 패턴을 재

결정하는 단계; 및

(g) 상기 트리의 노드에 맵핑된 이미지들 중 오류 이미지를 판별하고 판별된 오류 이미지와 대응하는 노드를 제거하는 단계를 더 포함하는,

이미지 캡차 제공 방법.

청구항 7

제1항에 있어서,

정답 또는 오답을 나타내는 상기 제1 그룹의 이미지의 선택에 의해 상기 이미지 캡차의 문제 통과 여부를 결정하는 단계; 및 통과된 경우 미리 결정된 행동 패턴을 이용하여 접속대상의 행동 패턴을 판별하는 단계;를 더 포함하는,

이미지 캡차 제공 방법.

청구항 8

제7항에 있어서,

상기 제2 그룹의 이미지는 정답 이미지 또는 오답 이미지의 분류 전의 수집된 이미지를 더 포함하며,

상기 이미지 캡차 제공 방법은, 접속대상이 사람의 행동 패턴을 가지는 것으로 판별된 경우, 상기 접속대상에 의한 상기 수집된 이미지에 대한 선택으로 정답 이미지 또는 오답 이미지로 분류하는 단계;를 더 포함하는,

이미지 캡차 제공 방법.

청구항 9

제8항에 있어서,

상기 단계 (a)는 반복적으로 수행되고,

상기 분류 단계 이후의 상기 단계 (a)의 복수의 이미지에 분류된 이미지가 제1 그룹의 이미지로 포함되는,

이미지 캡차 제공 방법.

청구항 10

제7항에 있어서,

상기 문제 통과 여부를 결정하는 단계는 상기 제2 그룹의 이미지에 대한 정답 또는 오답 선택에 상관없이 상기 제1 그룹의 이미지의 선택에 의해서만 상기 이미지 캡차의 문제 통과 여부를 결정하는,

이미지 캡차 제공 방법.

청구항 11

제1 그룹의 이미지와 제2 그룹의 이미지를 포함하는 복수의 이미지로 구성된 이미지 캡차를 출력하고 출력된 상기 이미지 캡차의 이미지 선택을 수신하는 인터페이스부;

상기 이미지 선택을 나타내는 로그 데이터를 저장하는 저장부; 및

상기 수신된 이미지 선택에 따라 상기 이미지 캡차의 문제 통과 여부를 결정하는 제어부를 포함하며,

상기 제1 그룹의 이미지는 접속대상의 이미지 캡차 문제의 통과 결정에 이용되고,

상기 제어부는 저장된 복수의 로그 데이터로 행동 패턴을 결정하고 상기 행동 패턴은 상기 제1 그룹의 이미지 및 상기 제2 그룹의 이미지 중 하나 이상의 이미지에 대한 선택을 나타내고,

상기 저장부에 저장되는 상기 행동 패턴은 루트 노드로부터 하나 이상의 자식 노드를 가지는 트리로 표현되는,

이미지 캡차 제공 서버.

청구항 12

삭제

청구항 13

제11항에 있어서,

상기 루트 노드로부터 지정된 리프 노드까지의 패스가 사람을 나타내는 행동 패턴으로 지정되며, 상기 루트 노드는 이미지에 맵핑되는,

이미지 캡차 제공 서버.

청구항 14

제11항에 있어서,

상기 제어부는 상기 인터페이스부를 통해 수신된 정답 또는 오답을 나타내는 상기 제1 그룹의 이미지의 선택에 의해 상기 이미지 캡차의 문제 통과 여부를 결정하고 통과된 경우 미리 결정된 행동 패턴을 이용하여 접속대상의 행동 패턴을 판별하는,

이미지 캡차 제공 서버.

청구항 15

제14항에 있어서,

상기 제2 그룹의 이미지는 정답 이미지 또는 오답 이미지의 분류 전의 수집된 이미지를 더 포함하며,

상기 제어부는, 접속대상이 사람의 행동 패턴을 가지는 것으로 판별된 경우, 상기 접속대상에 의한 상기 수집된 이미지에 대한 선택으로 정답 이미지 또는 오답 이미지로 분류하여 이후 분류된 이미지를 상기 복수의 이미지에 제1 그룹의 이미지로 포함하여 상기 인터페이스부를 통해 출력하는,

이미지 캡차 제공 서버.

발명의 설명

기술 분야

[0001] 본 발명은 이미지 캡차 제공 방법 및 이미지 캡차 제공 서버에 관한 것으로서, 구체적으로는 접속한 대상이 사람인지를 판별하는 데 이용되지 않는 이미지들을 제공하여 이미지의 신분 노출을 방지하고 접속대상의 패턴 분석에 따라 신규의 이미지를 이미지 캡차에서 활용 가능하도록 하는, 이미지 캡차 제공 방법 및 이미지 캡차 제공 서버에 관한 것이다.

배경 기술

[0002] 캡차(CAPTCHA: Completely Automated Public Turing test to tell Computers and Humans Apart)는 자동화된 튜링 테스트(Turing test)로서 원격에서 현재 접근하는 대상이 사람인지 아닌지를 판별하기 위해 만들어 졌다. 스팸메일을 배포하는 악성 유저들은 자동으로 정해진 작업을 수행하는 로봇을 활용하여 다수의 이메일(e-mail) 계정을 확보하려고 시도하는데 캡차는 로봇이 해결할 수 없고 사람은 해결할 수 있는 문제를 제시하여 이러한 작업을 저지한다. 현재 가장 많이 사용되는 곳은 구글, Bing 등의 대형 포털 서비스로서 이메일 계정 등의 리소스를 제공하는 과정에 사람임을 확인할 때 활용되고 있다.

[0003] 현재 널리 사용되고 있는 캡차는 성능과 관련하여 많은 논란이 일고 있다. 주로 사용되는 문자 기반 캡차는 그 취약함이 알려져 있고 캡차를 자동으로 해석하는 방법에 대해서도 또한 알려져 있다. 이러한 문제는 캡차가 사람만이 해결할 수 있는 문제를 제출해야 하지만 문자 기반 캡차의 핵심인 문자를 해석하는 능력에 있어 사람과 로봇이 그 차이가 없는 점에 기인한다. 문자 기반 캡차에 패턴 왜곡, 배경 그림 추가, 색 변경 등 다양한 기법이 적용되고 있지만 이는 오히려 사람에게도 문제를 풀기 힘든 상황을 야기한다. 오히려 로봇이 사람보다 더 문제를 쉽게 풀 수 있다는 지적이 나오고 있다.

[0004] 문자 기반 캡차의 문제점을 개선하기 위해서 이미지 기반 캡차가 등장하였다. 이미지 기반 캡차는 주로 이미지의 의미를 해석하는 문제로서 이미지 의미의 해석 능력이 사람과 로봇 간에 그 차이가 현격하기에 이상적인 캡

차를 구성할 수 있다.

[0005] 이미지의 의미 해석을 문제로 제출하는 초창기의 이미지 기반 캡차는 문제를 빨리 만들 수 없는 한계가 존재한다. 이는 이미지의 의미를 해석하는 작업이 사람만이 가능하기에 문제 제출 시에도 사람의 개입이 반드시 필요하기 때문이다. 초창기의 이미지 기반 캡차는 필연적으로 문제의 총 숫자가 적을 수밖에 없고 로봇의 컴퓨팅 파워를 이용한 문제 소모성 공격(이전에 틀렸던 경험을 활용하여 이미지의 정답 여부를 찾아내는 방식)을 통해 이미지 해석을 통하지 않고도 로봇은 이미지 기반 캡차의 통과가 가능하였다.

[0006] 문제 제출 속도의 한계를 극복하기 위해서 자동으로 문제를 만드는 방식이 또한 고안되었다. 이러한 후기의 이미지 기반 캡차는 주로 왜곡하거나 잘라낸 이미지에서 특정 패턴을 찾아내는 문제를 제출한다. 앞서 언급한 바와 같이 사람이 로봇에 비해서 우월한 능력을 가지는 것은 이미지를 해석하는 것이지 이미지의 패턴을 찾아내는 것이 아니기에 사람과 로봇을 판별하는 데 효과적이지 못하다.

[0007] 이와 같이 이미지의 의미 해석을 문제로 제출하는 초기의 이미지 기반 캡차의 한계를 극복하여 사람의 개입 없이 자동으로 문제의 숫자를 확장 가능하도록 하는 이미지 캡차 제공 방법 및 이미지 캡차 제공 서버가 필요하다.

발명의 내용

해결하려는 과제

[0008] 본 발명은 상술한 문제점을 해결하기 위해서 안출한 것으로서, 이미지 기반 캡차의 선택 패턴의 분석을 통해 사람의 개입 없이 자동으로 이미지 해석에 이용될 신규 이미지를 추가 가능하고 로봇에 의한 분석이 불가능하도록 하는, 이미지 캡차 제공 방법 및 이미지 캡차 제공 서버를 제공하는 데 그 목적이 있다.

과제의 해결 수단

[0009] 상기와 같은 목적을 달성하기 위한 이미지 캡차 제공 방법은 복수의 이미지로 구성된 이미지 캡차를 출력하는 단계 및 출력된 이미지 캡차의 이미지 선택을 수신하는 단계를 포함하며 복수의 이미지는 접속대상의 이미지 캡차 문제의 통과 결정에 이용되는 제1 그룹의 이미지 및 접속대상의 이미지 캡차 문제의 통과 결정에 이용되지 않는 제2 그룹의 이미지를 포함한다.

[0010] 또한 상기와 같은 목적을 달성하기 위한 이미지 캡차 제공 서버는 복수의 이미지로 구성된 이미지 캡차를 출력하고 출력된 이미지 캡차의 이미지 선택을 수신하는 인터페이스부 및 수신된 이미지 선택에 따라 이미지 캡차의 문제 통과 여부를 결정하는 제어부를 포함하며 복수의 이미지는 접속대상의 이미지 캡차 문제의 통과 결정에 이용되는 제1 그룹의 이미지 및 접속대상의 이미지 캡차 문제의 통과 결정에 이용되지 않는 제2 그룹의 이미지를 포함한다.

발명의 효과

[0011] 상기와 같은 본 발명에 따른 이미지 캡차 제공 방법 및 이미지 캡차 제공 서버는 이미지 기반 캡차의 선택 패턴의 분석을 통해 사람의 개입 없이 자동으로 이미지 해석에 이용될 신규 이미지를 추가 가능하고 로봇에 의한 분석이 불가능하도록 하는 효과가 있다.

도면의 간단한 설명

[0012] 도 1은 이미지 캡차 제공을 위한 시스템의 예를 도시한 도면이다.

도 2는 이미지 캡차 제공 서버의 예시적인 블록도를 도시한 도면이다.

도 3은 이미지 캡차 제공을 위한 예시적인 제어 흐름을 도시한 도면이다.

도 4는 수집된 이미지의 신분을 결정하고 행동 패턴의 재결정을 위한 예시적인 흐름을 도시한 도면이다.

도 5는 접속대상 장치로 출력되어 접속대상 장치에서 디스플레이되는 이미지 캡차의 예를 도시한 도면이다.

도 6은 기존 방식의 이미지 캡차의 문제 통과 방식의 예를 나타내는 도면이다.

도 7은 본 발명에 따른 이미지 캡차의 문제 통과 방식의 예를 나타내는 도면이다.

도 8은 행동 패턴들을 클러스터링 기법을 통해 분류하는 예를 도시한 도면이다.

도 9는 트리의 재구성에서 자동적으로 판별 가능한 전형적인 오류 패턴을 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0013] 상술한 목적, 특징 및 장점은 첨부된 도면을 참조하여 상세하게 후술 되어 있는 상세한 설명을 통하여 더욱 명확해 질 것이며, 그에 따라 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 또한, 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에 그 상세한 설명을 생략하기로 한다. 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시 예를 상세히 설명하기로 한다.
- [0014] 도 1은 이미지 캡차 제공을 위한 시스템의 예를 도시한 도면이다.
- [0015] 도 1에 따르면 이미지 캡차 제공 시스템은 적어도 하나 이상의 접속대상 장치(200) 및 이미지 캡차 제공 서버(100)를 포함한다. 접속대상 장치(200) 및 이미지 캡차 제공 서버(100)는 바람직하게는 인터넷망을 통해서 연결되어 각종 데이터를 송수신할 수 있도록 구성된다.
- [0016] 이미지 캡차 제공 시스템의 예를 통해 각 시스템 블록을 간단히 살펴보면, 접속대상 장치(200)는 인터넷망을 통해 서버나 다른 장치에 접속하고 인터넷망을 통해 각종 데이터의 송수신이 가능한 장치이다. 접속대상 장치(200)는 예를 들어 개인용 컴퓨터, 휴대폰, 스마트폰, 태블릿 PC, 특정용도의 서버 등일 수 있다. 접속대상 장치(200)는 적어도 이미지 캡차 제공 서버(100)에 접속할 수 있고 이미지 캡차 제공 서버(100)에서 제공하는 특정 기능을 이용할 수 있는 장치이다.
- [0017] 접속대상 장치(200)는 예를 들어 이미지 캡차 제공 서버(100)에 신규로 가입하기 위해 특정 개인(사람)이 이용하는 장치이거나 이미지 캡차 제공 서버(100)의 기능을 악의적으로 활용하기 위해 탑재된 프로그램을 통해 이미지 캡차 제공 서버(100)에 가입하기 위한 소위 로봇일 수 있다. 악의적인 의도의 로봇은 문제 소모성 공격을 통해 이미지 캡차의 이미지들을 분석하고 분석을 통해 다수의 계정을 획득할 수 있도록 프로그래밍된다.
- [0018] 이미지 캡차 제공 서버(100)는 가입 또는 등록된 사용자에게 서버의 용도에 따른 기능을 제공한다. 이미지 캡차 제공 서버(100)는 예를 들어 이메일 서버, 포털 서버, 클라우드 서버 등일 수 있다. 본 발명에 따른 이미지 캡차 제공 서버(100)는 이미지 캡차를 출력하고 출력된 이미지 캡차의 이미지들에 대한 선택으로 접속대상 장치(200)가 사람인지 아닌지를 판별할 수 있도록 구성된다. 이미지 캡차 제공 서버(100)는 이미지 캡차 문제를 통과한 경우 해당 사용자에게 계정을 제공하고 특정 기능이나 서비스를 제공할 수 있도록 구성된다. 또한 이미지 캡차 제공 서버(100)는 이미지 선택에 따른 행동 패턴으로 사람인지 여부를 판별할 수 있도록 구성된다.
- [0019] 접속대상 장치(200)와 이미지 캡차 제공 서버(100)는 인터넷망을 통해 인터넷 프로토콜상의 통신 프로토콜에 따라 각종 요청 및 응답 데이터를 송신할 수 있다. 접속대상 장치(200)는 계정 확보를 위한 요청을 이미지 캡차 제공 서버(100)로 송신하고 이미지 캡차 제공 서버(100)는 요청에 대한 응답으로서 복수의 이미지로 구성된 이미지 캡차를 제공한다. 제공된 이미지 캡차에 대한 접속대상 장치(200)에서의 선택과 선택을 나타내는 데이터 또는 응답의 수신으로 이미지 캡차 제공 서버(100)는 요청한 접속대상의 문제 통과 여부와 사람 여부를 판별하고 그 판별 결과에 따라 계정을 제공할 수 있다.
- [0020] 이미지 캡차 제공 서버(100)는 접속대상 장치(200)와 다양한 인터페이스를 가질 수 있다. 예를 들어 이미지 캡차 제공 서버(100)는 웹 페이지를 제공하는 웹 서버로서 동작할 수 있다. 이미지 캡차 제공 서버(100)는 웹 페이지를 접속대상 장치(200)로 전달하고 웹 페이지에 디스플레이된 이미지들 중에서의 선택에 따른 데이터를 수신할 수 있도록 구성된다.
- [0021] 본 발명에 따라 생성되는 이미지 캡차와 이 이미지 캡차를 활용하여 이루어지는 구체적인 제어 및 데이터 흐름은 이하의 도면을 통해서 상세히 살펴볼도록 한다.
- [0022] 도 2는 이미지 캡차 제공 서버(100)의 예시적인 블록도를 도시한 도면이다. 도 2에 따르면 이미지 캡차 제공 서버(100)는 인터페이스부(101), 저장부(103), 제어부(105) 및 연결부(107)를 적어도 포함하고 그외 다른 블록들을 더 포함할 수 있다. 도 2의 블록도는 바람직하게는 기능 블록도를 나타낸다.

- [0023] 도 2의 기능 블록도의 각 블록들을 간단히 살펴보면, 인터페이스부(101)는 인터넷망에 연결되어 하나 이상의 접속대상 장치(200) 각각과 데이터를 송수신한다. 인터페이스부(101)는 예를 들어 접속대상 장치(200)로 이미지 캡처를 나타내는 데이터를 송신하고 접속대상 장치(200)로부터 이미지 캡처에 포함된 이미지들에 대한 선택을 나타내는 데이터를 수신할 수 있다. 데이터 송수신을 위해 인터페이스부(101)는 유선 근거리 통신(예를 들어 이더넷 통신)이 가능한 칩셋이나 무선 근거리 통신(예를 들어 와이파이 통신)이 가능한 칩셋 등을 포함한다.
- [0024] 저장부(103)는 각종 데이터를 저장한다. 저장부(103)는 휘발성 메모리, 비휘발성 메모리 및/또는 하드디스크 등과 같은 대용량 저장매체를 적어도 하나 이상을 포함한다. 저장부(103)는 휘발성 메모리를 이용하여 각종 데이터와 프로그램을 임시로 저장하고 비휘발성 메모리를 활용하여 시스템 설정이나 시스템 프로그램(부팅 프로그램 등)을 저장한다. 저장부(103)는 하드디스크 등과 같은 대용량 저장매체를 이용하여 각종 이미지들과 각종 로그 데이터들과 각종 프로그램들을 저장한다.
- [0025] 저장부(103)에 저장되는 이미지들은 이미지 캡처의 구성이나 생성에 이용된다. 이 이미지들은 적어도 특정 문제에 대해서 정답 또는 오답인지가 결정된 이미지이거나 인터넷망을 통해 수집되었으나 특정 문제에 대해서 정답 또는 오답으로 아직 결정되지 않은 이미지일 수 있다. 이미지들은 이미지 캡처에 포함될 수 있다. 저장부(103)에 저장되는 프로그램들은 서버의 용도에 따른 기능을 제공하기 위한 프로그램, 이미지 캡처 생성과 이미지 관리 및 로그 데이터 관리를 위한 프로그램을 포함한다. 이미지 캡처 생성에 관련된 프로그램은 예를 들어 계정 생성시 호출되고 호출에 따라 이미지 캡처 문제가 접속대상 장치(200)로 전송할 수 있도록 구성된다.
- [0026] 제어부(105)는 이미지 캡처 제공 서버(100)의 각 블록들을 제어한다. 제어부(105)는 프로그램의 코드를 실행할 수 있는 실행 유닛을 포함한다. 제어부(105)는 바람직하게는 저장부(103)에 저장되어 있는 프로그램의 코드를 로딩하여 프로그램에 따라 각 블록들을 제어하도록 구성된다. 제어부(105)는 소위 하나 이상의 CPU, MPU 등을 포함하도록 구성된다.
- [0027] 제어부(105)는 인터페이스부(101)와 연동하여 각종 데이터를 접속대상 장치(200)로 송신하거나 수신할 수 있다. 제어부(105)는 이미지 캡처 제공 서버(100)에서 할당된 기능을 수행할 수 있고 접속대상 장치(200)로부터의 요청에 따라 해당 기능을 수행할 수 있도록 구성된다. 이 과정에서 제어부(105)는 접속대상 장치(200)가 사람인지를 판별할 필요가 있다. 예를 들어 제어부(105)는 신규 계정 생성시 접속대상 장치(200)가 사람인지를 판별하기 위해서 이미지 캡처를 생성하고 생성된 이미지 캡처를 인터페이스부(101)를 통해 출력한다. 제어부(105)는 이미지 캡처에 대한 접속대상 장치(200)의 선택을 인터페이스부(101)를 통해 수신하고 수신된 선택에 따라 접속대상 장치(200)가 사람인지를 판별하고 그 판별 결과에 따라 신규 계정을 접속대상 장치(200)에 제공한다.
- [0028] 제어부(105)에서 수행되는 제어 흐름은 도 3 및 도 4를 통해서 상세히 살펴보도록 한다.
- [0029] 연결부(107)는 블록 간의 데이터를 송수신한다. 연결부(107)는 제어부(105)에 의한 제어로 다른 블록의 데이터를 제어부(105)로 송신하거나 제어부(105)에 의해서 생성된 데이터를 다른 블록으로 송신하도록 구성된다. 연결부(107)는 예를 들어 병렬 버스로 구성된다.
- [0030] 도 2의 이미지 캡처 제공 서버(100)의 각 블록들은 단일의 기구물 내에 내장되거나 복수의 기구물 내에 분산 내장될 수 있다. 심지어 제어부(105)는 복수의 기구물 내에 분산 내장될 수도 있다. 이미지 캡처 제공 서버(100)의 구조에 따라 연결부(107)의 형태는 상이할 수 있다.
- [0031] 도 3은 이미지 캡처 제공에 이용될 행동 패턴 결정을 위한 예시적인 제어 흐름을 도시한 도면이다. 도 3의 제어 흐름은 이미지 캡처 제공 서버(100)에 의해서 수행된다. 바람직하게는 이미지 캡처 제공 서버(100)의 제어부(105)가 저장부(103)에 저장되어 있는 프로그램을 로딩하여 프로그램의 코드에 따라 인터페이스부(101) 및 저장부(103)를 제어하여 본 제어 흐름이 수행된다. 여기에서 각 블록들 간에 연동되는 제어 및 데이터 흐름을 상세히 살펴보도록 한다. 결정된 행동 패턴은 이후의 이미지 캡처 제공에 이용된다.
- [0032] 먼저 이미지 캡처 제공 서버(100)는 이 서버(100)의 관리자 등의 개입에 의해 씨드(Seed) 이미지들 및 이미지 캡처의 문제를 구성(S101)한다. 예를 들어 관리자 등은 이미지 캡처의 문제를 생성하고 씨드로 사용될 이미지들을 온라인/오프라인 등을 통해 수집하여 저장부(103)에 저장하고 생성된 문제에 대해서 각 씨드 이미지들이 정답 이미지인지 또는 오답 이미지인지를 분류한다. 단계 S101은 사람, 즉 관리자,에 의한 관여로 각 씨드 이미지의 의미 해석을 통해 각 씨드 이미지들의 정답 또는 오답인지가 결정되고 분류된다.
- [0033] 이후 이미지 캡처 제공 서버(100)의 제어부(105)는 인터페이스부(101)를 통한 접속대상 장치(200)의 (사용자)

계정 생성 요청에 따라 이미지 캡차를 구성(S103)한다. 구성되는 이미지 캡차는 복수의 이미지들을 포함하고 복수의 이미지들에 대한 선택 기준이 되는 문제를 더 포함한다. 복수의 이미지들은 바람직하게는 분류된 씨드 이미지들에서 선택된다.

[0034] 먼저 본 발명에 따른 이미지 캡차의 복수의 이미지들 모두는 정답 또는 오답 이미지로의 신분 인식이 불가능하도록 구성된다. 즉 구성된 이미지 캡차는 접속대상 장치(200)의 문제의 통과 여부 결정에 이용되는 그룹(이하 '제1 그룹'이라고도 함)의 이미지들과 접속대상 장치(200)의 문제 통과 결정에 이용되지 않는 그룹의 이미지들(이하 '제2 그룹'이라고도 함)을 포함한다. 제2 그룹의 이미지를 이하 '허니팟 이미지'라고도 지칭한다.

[0035] 알려진 이미지 캡차와는 달리 본 발명에 따른 이미지 캡차는 적어도 제1 그룹의 이미지들과 제2 그룹의 이미지들을 모두 포함하고 있어 설령 통과된 경우라도 각 이미지의 신분을 숨길 수 있도록 한다. 본 발명에 따른 이미지 캡차의 허니팟 이미지는 정답 또는 오답의 신분으로 참여하지 않는다. 허니팟 이미지는 이미 분류된 오답 이미지일 수도 있고 정답 이미지일 수도 있다. 설령 접속대상 장치(200)가 이미지 캡차 문제를 통과하더라도 허니팟 이미지에 의해 각 이미지들의 정답 또는 오답 신분을 결정할 수 없도록 한다. 허니팟 이미지에 대한 선택과는 상관없이 복수의 이미지들 중 일부의 이미지(제1 그룹의 이미지)에 의해서 통과 여부가 결정되기에 이미지들의 신분 확정이 불가능하고 문제 소모성 공격을 방지할 수 있다.

[0036] 즉 허니팟 이미지가 이미지 캡차에 포함됨에 따라, 접속대상 장치(200)인 로봇은 선택한 이미지 중에서 허니팟 이미지에 대한 잘못 선택으로도 문제를 통과할 수 있다. 허니팟 이미지는 이와 같이 정답 이미지이거나 오답 이미지일 수 있으나 문제 통과에 영향을 미치지 않는다. 로봇은 허니팟 이미지에 대한 정답 또는 오답으로의 착각으로 이후 해당 이미지(이전에는 허니팟 이미지이지만 이후에는 제1 그룹의 이미지로 선택될 수 있음)에 대한 동일한 선택을 할 수 있고 이에 따라 로봇을 방어하는 자료로 활용될 수 있다.

[0037] 허니팟 이미지는 이미지의 신분의 유추 시도를 교란시키는 역할과 선택 여부가 결과에 영향을 미치지 않음에 따라 아래에서 살펴볼 패턴 왜곡을 방지하고 또한 신규로 수집된 이미지에 대한 신분 결정에 이용된다.

[0038] 이미지 캡차의 구성 이후에 제어부(105)는 이미지 캡차를 인터페이스부(101)를 통해 접속대상 장치(200)로 출력(S105)한다. 출력된 이미지 캡차는 문제 및 복수의 이미지를 포함하고 복수의 이미지는 제1 그룹의 이미지(들) 및 제2 그룹(허니팟) 이미지(들)를 포함한다.

[0039] 도 5는 접속대상 장치(200)로 출력되어 접속대상 장치(200)에서 디스플레이되는 이미지 캡차의 예를 도시한 도면이다. 도 5에서 알 수 있는 바와 같이 접속대상 장치(200)는 인터넷망을 통해 이미지 캡차를 수신하고 수신된 이미지 캡차를 디스플레이한다. 접속대상 장치(200)의 사용자 또는 접속대상 장치(200)의 로봇은 이미지 캡차에 포함된 복수의 이미지들을 문제에 따라 선택하고 이미지 선택을 나타내는 데이터를 인터넷망을 통해 이미지 캡차 제공 서버(100)로 전송한다.

[0040] 이미지 선택을 나타내는 데이터를 인터페이스부(101)를 통해서 제어부(105)가 수신(S107)한다. 제어부(105)는 이미지 선택의 수신에 따라 이미지 캡차 문제의 통과 여부를 결정한다. 그 결정에 따라 제어부(105)는 계정을 생성하고 계정을 접속대상 장치(200)에 제공하거나 계정 생성을 거부한다.

[0041] 또한 이미지 선택을 나타내는 데이터는 로그 데이터로 저장부(103)에 제어부(105)에 의해 저장(S109)된다. 로그 데이터로 저장되는 데이터는 문제를 나타내는 데이터, 복수의 이미지의 식별자, 각 식별자의 이미지에 대한 선택 여부, 접속대상 장치(200)의 IP(Internet Protocol) 주소, 이미지 캡차 문제에 대한 통과 여부 등을 포함한다. 로그 데이터는 IP 주소별로 수집된다. IP 주소는 특정 접속대상 장치(200)를 식별할 수 있는 예시적인 식별자이다. 특정 접속대상 장치(200)는 필요에 따라 여러 번 본 이미지 캡차 제공 서버(100)에 접속할 수 있고 이에 따라 다수의 이미지 캡차 출력 및 수신 이력이 로그 데이터에 저장될 수 있다. 통과 여부의 데이터는 각 이미지 캡차 출력별로 저장될 수 있고 이로부터 통과 확률이 로그 데이터로 또한 저장될 수 있다.

[0042] 도 6은 기존 방식의 이미지 캡차의 문제 통과 방식을 나타내고 도 7은 본 발명에 따른 이미지 캡차의 문제 통과 방식을 나타낸다.

[0043] 도 6에서 알 수 있는 바와 같이, 기존의 알려진 이미지 캡차의 문제 통과 방식은 모든 정답 이미지를 선택하고 모든 오답 이미지를 선택하지 않으면 통과된다. 만일 통과된 경우에는 출력된 모든 이미지의 신분은 접속대상 장치(200)로 노출된다.

[0044] 반면에 도 7에서 알 수 있는 바와 같이, 제1 그룹의 이미지들은 정답 또는 오답 이미지의 선택으로 문제 통과 여부를 위해서 이용되고 제2 그룹(허니팟)의 이미지들은 문제 통과 결정에 이용되지 않는다. 따라서 만일 허니

팟 이미지를 잘못 선택하는 경우라도 접속대상 장치(200)는 통과할 수 있고 그 통과 결과로 이미지들의 신분 노출이 불가능하다. 제1 그룹의 이미지들의 개수는 예를 들어 14장 이상이다. 14장의 제1 그룹의 이미지들로 로복이나 랜덤(random) 알고리즘에 의한 우연한 통과를 방지할 수 있다.

[0045] 단계 S103 내지 단계 S109는 반복적으로 수행된다. 단계 S103 내지 단계 S109를 통해 출력되는 이미지 캡처의 이미지들은 이전 이미지 캡처의 이미지들과는 상이할 수 있다. 또한 이전 이미지 캡처의 제1 그룹의 이미지들은 이후 출력에서 허니팟 이미지로 출력될 수 있고 반대로 허니팟 이미지가 이후 출력에서 제1 그룹의 이미지로 출력될 수 있다. 충분한 로그 데이터가 저장부(103)에 저장되면 제어부(105)는 로그 데이터를 활용하여 특정 유형의 접속대상 장치(200) 또는 사람의 행동 패턴을 결정(S111)한다.

[0046] 행동 패턴의 결정, 특히 사람의 행동 패턴의 결정,은 수집된 이미지를 추가하는 것을 가능하도록 한다. 만일 특정 접속대상 장치(200)가 문제를 통과하고 나아가 사람의 행동 패턴과 동일한 것으로 판별되면, 수집된 미지의 이미지의 신분을 결정할 수 있고 이에 따라 수집된 이미지를 이미지 캡처에 새로운 이미지로 활용할 수 있다.

[0047] 행동 패턴의 결정은 크게 두개의 단계로 이루어진다. 첫 번째 단계는 저장된 로그 데이터를 이용하여 공통된 모습을 보여주는 패턴들을 클러스터링(clustering) 기법을 통해 분류하는 것이고 두 번째 단계는 사람에 의한 도움 없이 분류된 패턴들 중 하나를 사람의 행동 패턴으로 지정하는 것이다. 두개의 단계는 저장부(103)에 저장된 로그 데이터를 이용하여 제어부(105)에 의해서 수행되고 바람직하게는 이미지 선택 수신(S107)에 따라 이미지 캡처 문제를 통과한 로그 데이터의 이미지들을 대상으로 하여 공통된 패턴들을 클러스터링한다. 행동 패턴은 선택 패턴으로 이하 지칭될 수도 있고 행동 패턴은 바람직하게는 적어도 분류된 접속대상 장치(200)들에 의해 선택된 이미지들의 공통된 선택 패턴을 나타낸다.

[0048] 도 8은 첫 번째 단계에 대한 예를 도시하고 있다. 도 8에서 알 수 있는 바와 같이, 제어부(105)는 로그 데이터의 IP 주소와 사용된 이미지들을 이용하여 모든 IP 주소(도 8에서는 7개)를 두 그룹으로 나눌 수 있는 루트 노드의 이미지(도 8의 (a) 참조)를 찾아낸다. 왼쪽에 속한 4개의 IP는 루트 노드의 이미지를 정답으로 선택한 IP이고 오른쪽의 3개의 IP는 오답으로 선택한 IP들이다. 또한 이 두개의 그룹은 각 그룹 내에서 더 이상 다르게 선택한 IP가 존재하지 않을 때까지 두 그룹으로 나눌 수 있는 이미지를 반복(재귀)적으로 선택하고 트리를 구성한다. 도 8의 (b) 트리(tree)의 4개의 그룹(s3,s4,s5,s6)은 문제에 등장한 모든 이미지에 대해서 똑같이 대답하였기 때문에 더 이상 트리를 나눌 수 없는 상태이다.

[0049] 클러스터링 기법을 통한 분류에서 트리를 활용하고 있는 데, 트리를 활용하는 이유는 사람의 행동 패턴을 판별하기 위해서 필요한 이미지 숫자를 줄이기 위함이다. 사람의 행동 패턴을 트리 형태로 학습하면 사람과 그의 행동 패턴의 그룹을 나누는 것은 루트 노드로부터 사람 그룹까지의 패스(Path)에 존재하는 이미지만을 문제에 제출하는 것으로 충분하다. 예를 들어 s5 그룹이 사람으로 판별되면 s5 그룹까지의 패스에 존재하는 이미지를 문제에 포함하는 것으로 사람과 비사람 그룹간의 구별이 가능하고 이러한 장점은 그룹의 숫자가 증가할수록 트리의 깊이(depth)가 깊어질수록 유용하다.

[0050] 트리를 활용할 때의 장점을 극대화하기 위해서는 트리의 깊이를 줄이는 것이 필요하고 루트 또는 다른 노드의 좌우가 균형 잡힌 트리를 만드는 것이 중요하다. 이에 따라 본 발명에 따른 트리는 ID3 디시전 트리(decision tree)에서 사용되는 엔트로피를 이용하여 루트 노드 또는 자식 노드의 이미지 선정에 활용하였다.

[0051] 기본적으로 알려진 엔트로피 수식(http://en.wikipedia.org/wiki/ID3_algorithm 참조)은

수학식 1

$$H(S)=-\sum_{x \in X} p(x) \log_2 p(x)$$

[0052]

[0053] 이다. 여기서 S는 현재 엔트로피가 계산되는 데이터 세트를 나타내고 X는 S상의 클래스의 세트를 나타내고 p(x)는 세트 S상에서의 요소(elements)의 개수에 대한 클래스 x에서의 요소 개수의 비율을 나타낸다. 엔트로피 수식은 ID3 알고리즘에 따라 루트 노드로부터 자식노드로 반복적으로(재귀적으로) 수행되고 각 노드에 대해 특정 하나의 이미지가 선택될 수 있다.

[0054] ID3 트리와 관련하여 기존에 알려진 수학식 1에 따른 엔트로피 수식은 본 발명에 적용하기 위해서 변형이 필요하다. 만약 이미지 A가 문제(또는 시험)에 자주 제출되어 모든 IP에 대한 로그 데이터가 존재하지만 이미지 B는

문제에 거의 출제되지 않아 대부분의 IP에 대한 로그 데이터가 존재하지 않는다면 아무리 이미지 B의 엔트로피 수치가 높게 나온다 하더라도 경험하지 못한 IP를 둘로 나누는데 도움이 되지 않는다. 따라서 실제 트리 생성시 두 그룹으로 균등하게 나누는 능력을 보여주는 기준 공식에 얼마나 많은 IP를 나눌 수 있는지를 반영한 변형 공식을 본 발명에서는 이용한다. 변형된 엔트로피 수식은

수학식 2

$$H(S) = -\frac{k}{T} \sum_{x \in X} p(x) \log_2 p(x)$$

[0055]

와 같다. 여기서 T는 나누려하는 IP 그룹의 모든(총) IP 개수를 나타내고 k는 나누려 하는 IP 그룹에서의 이미지 S를 경험했던 IP의 총 개수를 나타낸다.

[0056]

수학식 2에 따라, 트리의 루트 노드에 맵핑된 이미지는 로그 데이터를 통해 식별되는 모든 IP의 개수에 대하여 해당 이미지 S를 경험한 총 IP 개수를 이용하여 가장 높은 엔트로피를 가지는 이미지로 선정된다. 루트 노드 이하의 자식 노드 또한 각 노드에 대해서 클러스터링된 IP 그룹의 모든 IP의 개수에 대하여 특정 이미지 S를 경험한 총 IP 개수를 이용하여 그 그룹 내에서 가장 높은 엔트로피를 가지는 이미지를 선정하여 맵핑된다.

[0057]

도 8의 (b)는 이러한 트리 구성과 이미지 맵핑에 따라 구성되는 트리의 예를 나타낸다. 도 8의 (b)에서 알 수 있는 바와 같이 각 그룹별 행동 패턴은 루트 노드로부터 자식 노드를 가지는 트리로 표현된다. 각 그룹별 행동 패턴은 단계 S107을 통해서 수신된 제1 그룹 및 허니팟(제2 그룹) 이미지(들)의 선택을 나타내고 바람직하게는 선택에 따른 클러스터링 분류를 나타낸다.

[0058]

트리의 구성 이후에 제어부(105)는 트리의 행동 패턴별로 클러스터링되어 있는 리프 노드 중 특정 하나의 리프 노드를 사람의 행동 패턴으로 지정한다. 제어부(105)는 간단하게 특정 하나의 리프 노드를 사람을 나타내는 행동 패턴으로 지정할 수 있다. 제어부(105)는 로그 데이터에 포함되어 있는 통과 여부 또는 통과 확률이 가장 높은 IP 그룹의 리프 노드를 사람을 나타내는 행동 패턴으로 지정한다.

[0059]

이후 제어부(105)는 구성된 트리에 동일한 분류 기능을 수행할 수 있는 이미지를 추가(S113)한다. 제어부(105)는 적어도 루트 노드로부터 사람을 나타내는 행동 패턴의 리프 노드까지의 패스상의 노드에 대해서 이미지를 추가하여 사람의 행동 패턴의 리프 노드 이전의 노드 각각에 대해서 다수의 이미지를 확보한다.

[0060]

예를 들어 도 8의 (b)와 같은 트리가 존재한다면 트리의 특성을 이용하여 간단하게 이미지를 특정 노드에 맵핑하여 추가할 수 있다. 도 8의 (b)의 왼쪽 자식 노드의 이미지는 (s3)와 (s4)를 구분하는 목적을 가지고 있는 이미지이다. 즉 (s3)에 속하는 모든 IP가 모두 선택(pick)하고 동시에 (s4)의 모든 IP가 공통적으로 선택하지 않은 이미지 또는 반대로 (s3)에 속하는 모든 IP는 모두 선택하지 않고 (s4)의 모든 IP가 모두 선택한 이미지는 동일한 기능을 수행할 수 있다. 루트 노드에도 또한 동일하게 새로운 이미지가 추가될 수 있다.

[0061]

트리의 루트 노드로부터 사람의 행동 패턴을 나타내는 리프 노드까지 이전의 각 노드에 맵핑된 이미지들은 정답 또는 오답의 선택과 그에 따른 통과 여부를 결정하는 제1 그룹의 이미지들로부터 선택되거나 통과 여부에 상관없는 제2 그룹(허니팟) 이미지들로부터 선택된다. 따라서 사람의 행동 패턴의 식별은 통과 여부와는 상관없이 출력된 이미지의 선택으로 결정된다.

[0062]

사람의 행동 패턴을 판별하기 위한 트리는 저장부(103)에 저장되고 이후에 사람임을 판별하고 신규 이미지 추가에 활용될 수 있다.

[0063]

도 3의 제어 흐름은 적어도 사람의 행동 패턴을 판별할 수 있고 바람직하게는 사람의 행동 패턴이 트리로 표현된다. 사람의 행동 패턴은 이후 온라인/오프라인을 통해서 수집된 신규의 이미지의 신분을 결정하기 위해서 이용된다.

[0064]

도 4는 수집된 이미지의 신분을 결정하고 행동 패턴 재결정을 위한 예시적인 흐름을 도시한 도면이다. 도 4의 제어 흐름은 일반적으로 도 3의 수행 이후에 수행된다. 도 4의 제어 흐름은 이미지 캡처 제공 서버(100)에 의해서 수행된다. 바람직하게는 이미지 캡처 제공 서버(100)의 제어부(105)가 저장부(103)에 저장되어 있는 프로그램을 로딩하여 프로그램의 코드에 따라 인터페이스부(101) 및 저장부(103)를 제어하여 본 제어 흐름이

[0065]

수행된다. 여기에서 각 블록들과 연동되는 제어 및 데이터 흐름을 상세히 살펴보도록 한다. 결정된 행동 패턴은 이후의 이미지 캡차 제공에 이용된다. 도 4의 제어 흐름의 설명에서 도 3의 설명과 중복되는 부분은 간단하게 살펴보도록 한다.

- [0066] 먼저 제어부(105)는 주기적으로 온라인 또는 오프라인을 통해서 신규 이미지들을 수집하고 수집된 이미지들을 저장부(103)에 저장(S201)한다. 단계 S201은 다른 단계들과 독립적으로 수행될 수 있다.
- [0067] 제어부(105)는 도 3을 통해서 미리 결정된 사람의 행동 패턴을 이용하여 인터페이스부(101)를 통한 접속대상 장치(200)의 계정 생성 요청에 따라 이미지 캡차를 구성(S203)한다.
- [0068] 구성된 이미지 캡차는 문제 통과 여부를 결정하는 데 이용되는 제1 그룹의 이미지, 문제 통과 여부의 결정에 영향을 미치지 않는 허니팟 이미지를 포함한다. 이미지 캡차에 포함된 복수의 이미지들 중 사람의 행동 패턴을 나타내는 루트 노드로부터 리프 노드의 패스상의 각 노드에 맵핑된 이미지가 또한 각각 포함된다. 이미지 캡차의 이미지들은 사람의 행동 패턴을 나타내는 패스상의 각 노드의 이미지(들)를 적어도 하나 이상 포함하고 각각의 이미지는 제1 그룹의 이미지로 지정되거나 허니팟 이미지로 지정된다.
- [0069] 또한 이미지 캡차의 허니팟 이미지로 지정된 특정 하나 이상의 이미지는 온라인 또는 오프라인을 통해 수집된 이미지(단계 S201 참조)일 수 있다. 수집된 이미지에 대한 정답 또는 오답은 사람의 관여 없이는 알 수 없는 상태로 수집된 이미지는 정답 또는 오답 이미지의 분류 전의 이미지이다. 일반적으로는 이미지의 의미 해석을 통해서만 정답 또는 오답으로 분류가능하고 사람의 관여 없이는 그 분류가 불가능하다.
- [0070] 제어부(105)는 복수의 이미지 및 문제를 포함하는 이미지 캡차를 인터페이스부(101)를 경유하여 접속대상 장치(200)로 출력(S205)한다.
- [0071] 제어부(105)는 인터페이스부(101)를 통해 접속대상 장치(200)로부터 이미지 선택을 나타내는 데이터를 수신(S207)한다. 수신된 이미지 선택을 나타내는 데이터는 다른 데이터(예를 들어 IP 주소 등, 도 3의 S107 및 그의 설명 참조)와 함께 저장부(103)에 제어부(105)의 제어로 로그 데이터로 저장(S209)된다.
- [0072] 이후 제어부(105)는 수신된 이미지 선택(제1 그룹의 이미지의 선택)에 의해 이미지 캡차의 문제 통과 여부를 결정(S211)한다. 문제 통과 여부의 결정에 이용되는 이미지들은 제1 그룹의 이미지들만으로 국한된다. 바람직하게는 제2 그룹의 이미지들에 대한 정답 또는 오답의 선택에 상관없이 제1 그룹의 이미지들에 대한 정답 또는 오답의 선택으로 문제 통과 여부가 결정된다.
- [0073] 만일 이미지 캡차의 문제를 통과한 경우에, 제어부(105)는 접속대상 장치(200)의 행동 패턴을 판별(S213)한다. 이를 위해 제어부(105)는 저장부(103)에 저장되고 도 3 등을 통해 미리 결정된 사람의 행동 패턴과의 비교를 통해 접속대상 장치(200)의 이미지 선택에 따른 행동 패턴이 사람의 행동 패턴을 가지는 지를 판별한다.
- [0074] 구체적으로, 제어부(105)는 미리 저장되어 있는 트리로 사람의 행동 패턴을 식별할 수 있고 단계 S205를 통해 출력된 이미지(들)에 대하여 수신된 선택으로 저장부(103)에 저장된 행동 패턴의 트리의 루트 노드로부터 자식 노드로 경로를 찾아 특정 노드를 선택할 수 있다. 이 특정 노드가 사람의 행동 패턴을 나타내는 리프 노드인 경우에는 제어부(105)는 사람의 행동 패턴을 가지는 것으로 판별한다.
- [0075] 또한 제어부(105)는 단계 S211에서의 통과 결정에 따라 나아가 행동 패턴의 판별에 따라 신규 계정을 생성하고 생성된 계정을 접속대상 장치(200)로 인터페이스부(101)를 통해 통지할 수 있다.
- [0076] 제어부(105)는 접속대상 장치(200)의 판별된 행동 패턴에 따라 수집된 이미지를 분류(S215)한다. 구체적으로, 문제를 통과하고 접속대상 장치(200)의 이미지 선택에 따른 행동 패턴이 트리 형태로 미리 저장되어 있는 사람의 행동 패턴과(을) 동일한(가지는) 경우에, 제어부(105)는 미리 수집되어 허니팟 이미지로 출력된 이미지를 정답 또는 오답으로 분류한다. 즉, 제어부(105)는 수집된 이미지의 접속대상 장치(200)의 정답 또는 오답 선택의 수신에 따라 해당 수집된 이미지의 정답 또는 오답의 신분을 분류 및 결정할 수 있다. 오답 또는 정답의 분류는 복수 회에 걸친 이미지 선택의 수신(S207)로 결정될 수도 있다. 이에 따라 특정 수집된 이미지는 복수 회에 걸쳐서 사람의 패턴으로 판별된 접속대상 장치(200)로부터의 복수의 선택에 기초해서 그 신분이 결정된다.
- [0077] 단계 S203 내지 단계 S215는 접속대상 장치(200)의 요청에 따라 반복적으로 수행될 수 있다. 단계 S215에서 분류된 특정 수집된 이미지는 이후 반복에서 제1 그룹의 이미지에 포함되어 접속대상 장치(200)로 출력될 수 있다. 이러한 과정을 통해서 사람의 관여 없이도 수집된 이미지에 대한 문제의 정답 또는 오답을 사람의 패턴으로 판별된 접속대상 장치(200)의 선택으로 결정하고 신규의 이미지를 이미지 캡차에 추가할 수 있다. 이에 따라

이미지 해석을 통한 문제를 제출 가능하면서도 기존의 알려진 이미지 캡차의 문제 제출 속도를 개선할 수 있다.

- [0078] 단계 S203 내지 단계 S215의 반복에 따라 저장부(103)는 여러 접속대상 장치(200)로부터의 로그 데이터를 지속적으로 저장한다.
- [0079] 지속적으로 저장된 로그 데이터는 이후 학습을 통한 행동 패턴의 재결정에 이용될 수 있다. 일정한 기간의 경과 시에 및/또는 로그 데이터의 수집량이 지정된 데이터 양 이상이 경우 등에, 행동 패턴의 재결정이 수행될 수 있다.
- [0080] 일정 기간의 경과나 로그 데이터가 지정된 데이터 양 이상인 경우 등에, 제어부(105)는 저장부(103)에 저장된 로그 데이터를 이용하여 축적된 로그 데이터의 학습을 통해 행동 패턴을 재결정(S217)한다. 재결정을 하는 방식은 수학적 2를 이용하는 도 3의 S111과 동일하다. 재결정을 통해 접속대상 장치(200)인 로봇들의 변화하는 행동 패턴에 적응가능하고 출제된 다양한 이미지들에 대한 신규의 행동 패턴을 결정할 수 있다. 결정된 행동 패턴은 사람임을 나타내는 행동 패턴을 포함하고 사람을 나타내는 행동 패턴은 트리 상에서 표현되고 루트 노드로부터 특정 리프 노드까지의 패스로 표현된다.
- [0081] 사람으로 정확히 분류하기 위해서는 학습에 왜곡이 없는 행동(선택) 패턴을 사용하는 것이 중요하다. 도 4의 단계 S217, 나아가 도 3의 S111,에서 행동 패턴의 결정을 위한 학습에 이용되는 로그 데이터, 구체적으로 이미지들,는 이미지 캡차 문제를 통과한 경우의 로그 데이터(이미지들)이다. 본 발명은 허니팟 이미지를 활용하여 특정 접속대상 장치(200)의 패턴의 숨기려는 시도를 또한 차단한다.
- [0082] 접속대상 장치(200)인 로봇이 문제에 통과하기 위한 고유한 패턴을 가지고 있다. 만약 자신의 패턴을 숨기고자 한다면 자신의 알고리즘과 달리 행동해야 한다. 성공확률이 높은 알고리즘일수록 자신의 알고리즘과 달리 행동하면 문제(시험)에 실패할 확률이 높고 문제 통과가 어렵다. 만약 자신의 패턴을 숨기기 위해 임의로 이미지를 선택한다면 14장의 이미지만으로도 성공확률은 0.006%에 불과하다.
- [0083] 기존의 이미지 기반 캡차에 허니팟 이미지가 없었기 때문에 문제에 통과하는 경우만을 학습에 활용하면 모든 사람(사용자)과 로봇의 패턴이 똑같아지는 문제가 있다. 만약 문제에 통과한 경우만을 로그 데이터로 모은 경우 정답 이미지가 문제로 출제되었을 경우 선택받지 못하면 문제 통과가 불가능하여 선택받지 못한 로그 데이터는 제외되고 오답 이미지가 선택된 경우 문제 통과가 불가능하여 이 로그 데이터 또한 제외된다.
- [0084] 결국 성공한 경우만을 모은다면 문제에 대한 모든 결과 로그가 동일하여 행동 패턴의 학습이 불가능하다. 본 발명에 따라 제안된 허니팟 이미지는 문제 결과에 영향을 미치지 않고 이 특징을 활용하여 문제에 통과한 경우만을 행동 패턴의 학습에 활용할 수 있다. 문제에 통과한다 하더라도 정답 이미지이지만 선택받지 못한 허니팟 이미지와 오답 이미지이지만 선택받은 허니팟 이미지가 로그 데이터에 존재할 수 있다. 이미지 캡차 제공 서버(100)의 저장부(103)에 저장된 모든 이미지들은 허니팟 이미지로서 문제에 제출 가능하므로 충분한 횟수만 제공된다면 문제에 통과한 경우만의 로그 데이터의 이미지들을 이용하여 모든 이미지의 행동(선택) 패턴을 수집 가능하다.
- [0085] 허니팟 이미지는 결국 문제를 보는 모든 접속대상 장치(200) 또는 사용자에게 이미지의 정보를 숨기는 역할을 담당한다. 기존 이미지 캡차와 달리 허니팟 이미지가 포함된 문제는 문제 통과가 그 이미지들의 신분을 결정하지 않는다. 오히려 행동 패턴의 결정을 위해 학습된 트리를 이용하여 정답으로 선택된 오답 이미지를 빠르게 분류하여 이후 문제에서는 오답 이미지로 등장시킴으로써 시험에 통과가 불가능하다.
- [0086] 행동 패턴의 재결정을 위해서, 제어부(105)는 저장부(103)에 저장된 복수의 로그 데이터를 이용하여 행동 패턴을 나타내는 트리를 재구성한다. 트리의 재구성에 이용되는 알고리즘은 수학적 2를 활용한 도 3의 단계 S111과 동일하다. 재구성된 트리는 기존의 도 3에서 구성된 트리를 대체한다. 트리의 재구성에 이용되는 로그 데이터는 바람직하게는 단계 S211을 통해 문제를 통과한 경우의 이미지들이다. 트리의 재구성에 따라 제어부(105)는 사람을 나타내는 행동 패턴을 지정할 수 있다. 사람을 나타내는 행동 패턴은 루트 노드로부터 리프 노드상의 패스로 표현된다. 사람을 나타내는 행동 패턴 상의 이미지들은 이후 제1 그룹의 이미지로 또는 허니팟 이미지로 제출될 수 있다.
- [0087] 이후 제어부(105)는 도 3의 단계 S113과 동일하게 재구성된 트리에 동일한 분류 기능을 수행할 수 있는 이미지를 추가(S219)한다. 제어부(105)는 적어도 루트 노드로부터 사람을 나타내는 행동 패턴의 리프 노드 이전의 패스상의 노드에 대해서 이미지를 추가하여 사람의 행동 패턴의 리프 노드 이전의 노드 각각에 대해서 다수의 이미지를 확보한다.

- [0088] 이후 제어부(105)는 재구성된 트리의 노드에서 맵핑된 이미지들 중에서 오류 이미지를 제거(S221)한다. 이를 위해 제어부(105)는 트리의 맵핑된 이미지들 중 오류 이미지를 판별하고 판별된 오류 이미지와 대응하는 노드를 재구성된 트리에서 삭제한다. 행동 패턴의 재결정 과정에서 오류 이미지의 판별 및 삭제는 또한 도 3의 단계 S113 이후에서 동일하게 적용될 수 있다.
- [0089] 본 발명은 사람의 패턴을 판별하고 해당 패턴을 이미지 검증에 활용하나 사람은 언제라도 실수를 할 수 있기에 사람의 패턴을 분석한 결과가 항상 참일 수는 없다. 사람의 실수에 의한 전형적인 트리 패턴을 찾아서 오류 이미지를 배제할 필요가 있다.
- [0090] 도 9는 트리의 (재)구성에서 자동적으로 판별 가능한 전형적인 오류 패턴을 나타낸다.
- [0091] 오류가 발생하는 가장 빈번한 두 가지 원인은 사람에 의한 실수와 사람조차 판별하기 힘든 애매한 이미지이다. 도 9의 (a)는 실수 패턴으로서, 대표적인 특징은 (s1)과 (s2)에 속한 IP들은 비슷하게 높은 통과 확률(성공률)을 나타내고 루트 노드에는 동일한 기능을 수행할 수 있는 다수의 이미지가 포함된 반면 오른쪽 서브 트리는 소수의 이미지만 존재하고 해당 이미지가 나누고 있는 IP 그룹이 한쪽으로(예를 들어 왼쪽)으로 지나치게 기울어 있는 모습을 보여준다. 도 9의 (a)로부터 (s2)에 속한 IP는 (s1)과 대부분 행동 패턴이 동일하나 하나의 이미지에 한하여 다르게 행동해 자신과 동일하게 행동한 타 IP가 없음을 알 수 있다. 이러한 경우 오른쪽 서브 노드의 이미지(오류 이미지)만을 제거하고 (s1)과 (s2)를 하나로 결합하여 트리를 수정한다. 이에 따라 오류 이미지는 제거된다.
- [0092] 도 9의 (b)의 트리는 사람이 판별하기 애매한 이미지가 문제에 포함된 경우를 나타내는 전형적인 패턴이다. 도 9의 (a)와 유사하게 (s3)와 (s4)는 높은 수준의 통과 확률(성공률)을 보이고 있고 루트 노드에는 동일한 기능을 수행하는 다수의 이미지가 포함되나 오른쪽 서브 트리의 노드는 하나의 이미지만에 대응한다. 도 9의 (a)와 달리 애매한 이미지는 균등하게 IP들을 분류하고 있다. 도 9의 (b) 역시 해당 오류 이미지와 대응하는 노드를 (재)구성된 트리로부터 제거하고 (s3)와 (s4)를 결합하여 오류 이미지가 제거된 트리가 구성된다.
- [0093] 위 두 가지 두 가지 패턴의 특징은 항상 문제 노드의 이미지의 숫자가 타 노드에 비해서 적고 문제 통과 확률이 높다는 것이다. 예를 들어 제어부(105)는 실수 패턴과 애매한 이미지를 나타낼 수 있는 노드의 이미지의 개수가 임계치 이하이고 해당 노드의 통과 확률이 다른 노드의 통과 확률보다 높은 경우 해당 이미지 및 노드를 제거할 수 있다. 임계치는 지정된 개수로 미리 결정되거나 다른 노드의 이미지 개수에 따라 결정될 수 있다.
- [0094] 오류 이미지의 제거 후의 행동 패턴을 나타내는 트리는 이후 단계 S203 내지 단계 S215에서 사람을 나타내는 행동 패턴을 판별하기 위해서 이용된다. 구체적으로는 사람을 나타내는 행동 패턴을 나타내는 루트 노드로부터 리프 노드까지의 각 노드의 이미지들 중 적어도 하나 이상은 이미지 캡처에 포함되어 접속대상 장치(200)로 출력된다. 이 이미지는 제1 그룹의 이미지로 출력되거나 허니팟 이미지로 출력될 수 있다. 물론 제1 그룹의 이미지와 허니팟 이미지는 이미지 캡처에서의 위치가 고정되지 않으며 가변적으로 변경될 수 있고 출력되는 각 이미지 캡처별로 상이한 위치에 제1 그룹의 이미지 및 허니팟 이미지가 출력된다.
- [0095] 본 발명에 따른 이미지 캡처는 허니팟 이미지를 포함하여 문제를 통과하는 경우라도 각 이미지의 신분 노출이 불가능하도록 한다. 또한 본 발명은 허니팟 이미지에 따른 정답 또는 오답 분류의 불가능으로 로봇의 행동 패턴과 사람의 행동 패턴을 분류가능하다. 또한 본 발명은 사람의 행동 패턴으로 수집된 신규의 이미지를 문제로 제출할 수 있고 신규의 이미지에 대한 신분을 결정하고 이후 신규 수집된 이미지를 문제에 활용할 수 있도록 한다.
- [0096] 이상에서 설명한 본 발명은, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 있어 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 여러 가지 치환, 변형 및 변경이 가능하므로 전술한 실시 예 및 첨부된 도면에 의해 한정되는 것이 아니다.

부호의 설명

[0097] 100 : 이미지 캡처 제공 서버

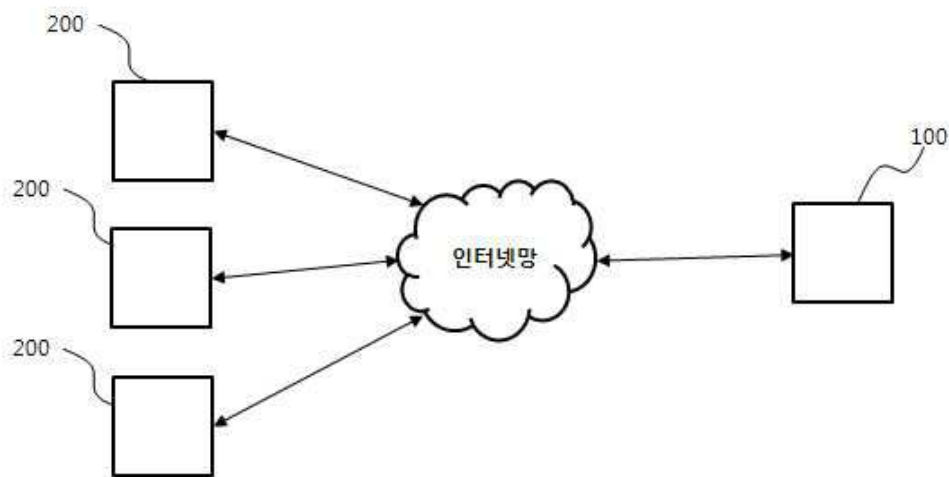
101 : 인터페이스부

103 : 저장부

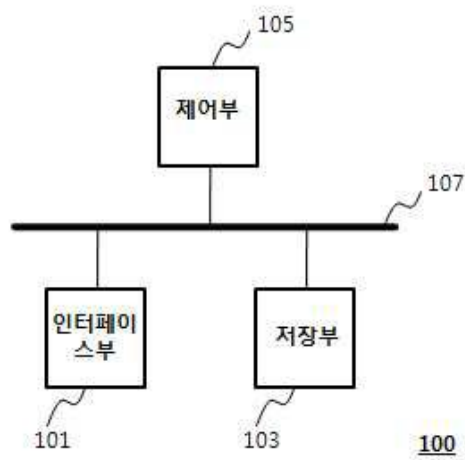
105 : 제어부
107 : 연결부
200 : 접속대상 장치

도면

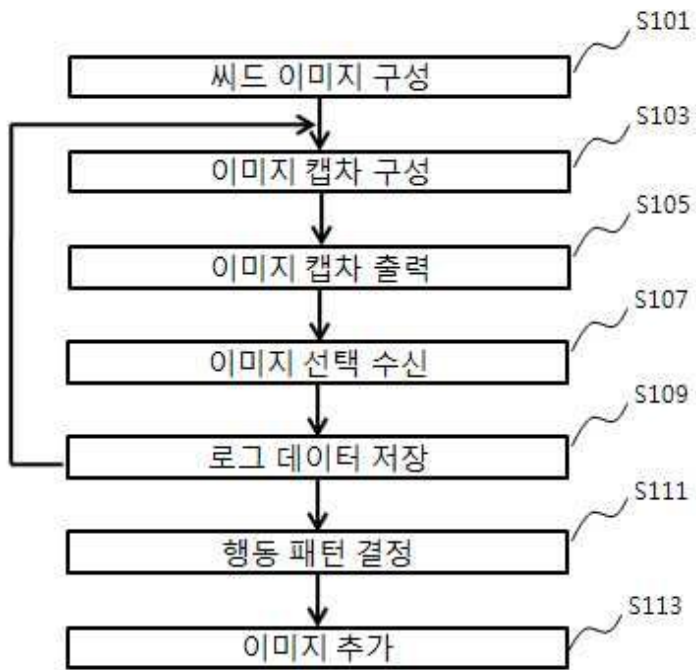
도면1



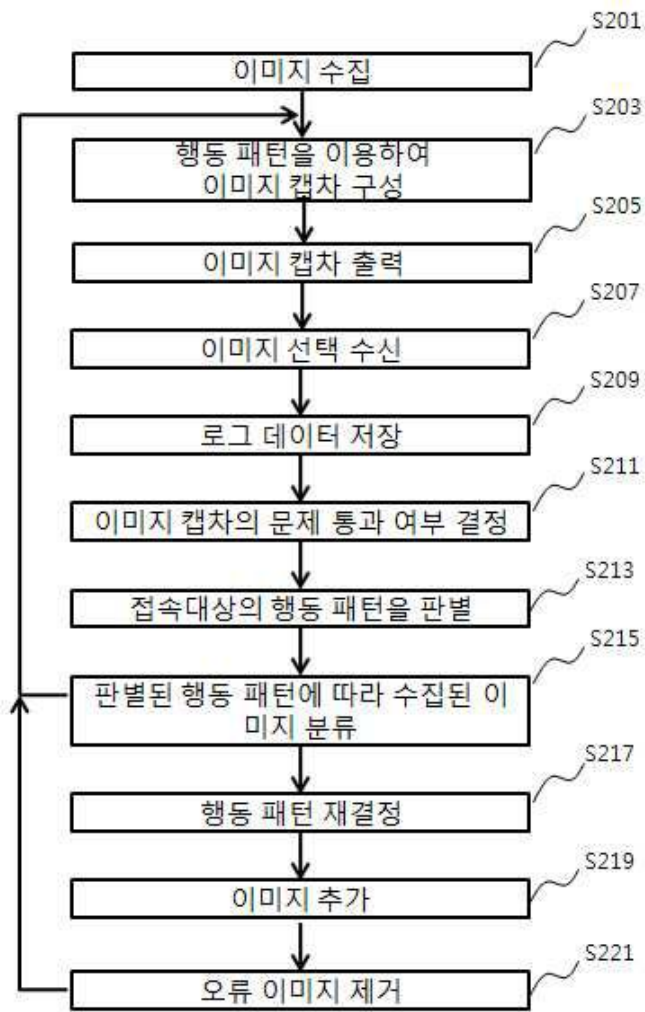
도면2



도면3



도면4



도면5



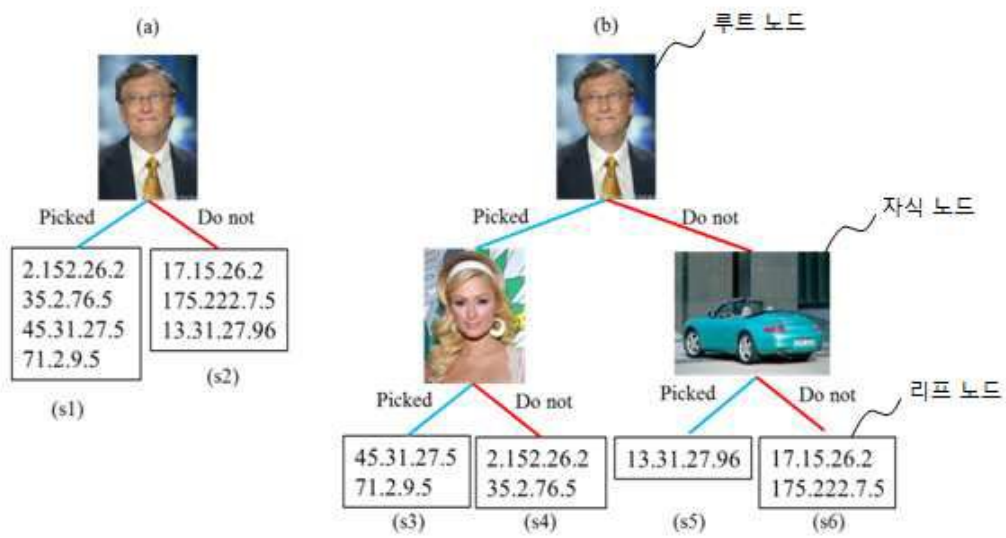
도면6



도면7



도면8



도면9

