

음성인식 유튜브 영상 조절 애플리케이션

소프트웨어 요구사항 명세서 (Software Requirement Specification)

2020년 4월 10일

인하대학교 컴퓨터공학과
캡스톤설계 - 001반

교 수 신병석
12163588 김예담
12130397 박찬준
12152984 전은호

목 차 (Table of Contents)

1. 서론 (Introduction)	4
1.1 목적 (Purpose)	4
1.2 대상 독자(Intended Audience)	4
1.3 범위 (Scope)	4
1.4 용어 및 약어의 정리 (Definitions, Acronyms, and Abbreviations)	4
1.6 개요 (Overview)	5
2. 종합 기술 (Overall Description)	6
2.1 프로젝트 결과물의 전망 (Product Perspective)	6
2.1.1 시스템 인터페이스 (System Interfaces)	6
2.1.2 사용자 인터페이스 (User Interfaces)	6
2.1.3 소프트웨어 인터페이스 (Software Interfaces)	6
2.1.4 운영 (Operation)	6
2.2 결과물의 기능 (Product Functions)	6
2.3 사용자의 특성 (User Characteristics)	7
2.4 제약 조건 (Constraints)	7
2.5 가정과 의존성 (Assumptions and Dependencies)	7
3. 상세 요구사항 (Specific Requirements)	7
3.1 외적 인터페이스 요구사항 (External Interface Requirements)	7
3.1.1 로딩화면	7
3.1.2 영상목록 화면	8
3.1.3 영상 재생화면	11
3.1.4 목록으로 이동하기 확인 대화상자	11
3.1.5 Keyword Spotting이 일어난 화면	12
3.1.6 음성을 통해 영상 멈춤	12
3.1.7 음성을 통해 영상 다시 재생	13
3.1.8 음성을 통해 영상 간격 이동	13
3.1.9 음성을 통해 북마크 생성	13
3.1.10 음성을 통해 북마크로 이동	13
3.1.11 음성을 통해 다음 영상 재생	13
3.1.12 음성을 통해 특정 영상 재생	13
3.1.13 음성을 통해 재생목록으로 이동	13
3.1.14 음성을 통해 프로그램 종료	14
3.2 기능적 요구사항 (Functional Requirements)	15
3.3 성능 요구사항 (Performance Requirements)	18
3.4 논리적 데이터베이스 요구사항 (Logical Database Requirements)	18
3.5 설계 제약사항 (Design Constraints)	18

3.6 소프트웨어 시스템 특성 (Software System Attributes)	18
3.7 기능별 구현 우선순위 (Functional Development Priority)	18
3.7.1 우선 구현사항	18
3.7.2 선택 구현사항	18

1. 서론 (Introduction)

1.1 목적 (Purpose)

본 문서는 음성인식 유튜브 영상 조절 애플리케이션 ‘쿠키야! - 핸드프리 비디오 플레이어’(이후 쿠키야!)를 위한 소프트웨어 요구사항 명세서(SRS)이다. 본 문서는 인하대학교 컴퓨터공학과 ‘컴퓨터공학 종합설계’ 001분반 ‘예찬호’ 팀의 프로젝트 설계 및 구현을 위한 것으로 이를 위한 요구 사항을 정리, 분석하고, 기재된 내용을 바탕으로 시스템을 설계 및 구현한다.

1.2 대상 독자(Intended Audience)

본 문서는 인하대학교 ‘컴퓨터공학 종합설계’ 001분반 ‘예찬호’ 팀을 주요 독자로 한다. 부가적으로 담당 교수 신병석 교수님이 주요 독자가 될 수 있다. 추후 본 시스템을 상품으로 개발할 경우, 이와 관련된 모든 업체 직원들이 추가적인 독자가 될 수 있다.

1.3 범위 (Scope)

본 프로젝트는 음성을 통해서 유튜브 영상을 컨트롤하는 새로운 애플리케이션을 개발하는 것이다. 본 프로젝트의 개발 범위는 다음과 같다.

- 특정 단어가 발화되었는지 인식할 수 있는 Keyword Spotting 기능
- 음성을 통해서 영상을 멈추고, 다시 재생하는 기능
- 음성을 통해서 영상의 플레이 시점을 조절하는 기능
- 음성을 통해서 영상의 특정 시간대에 북마크를 설정하고, 특정 북마크로 이동하는 기능
- 음성을 통해 컨트롤할 유튜브 영상을 애플리케이션 내부로 저장하는 기능
- 음성을 통해서 ‘쿠키야!’ 애플리케이션을 동작하는 기능

1.4 용어 및 약어의 정리 (Definitions, Acronyms, and Abbreviations)

- **Activity, 액티비티** : 안드로이드 애플리케이션의 핵심 구성요소 중의 하나로서 사용자에게 보여지는 하나의 화면을 말한다.
- **Android, 안드로이드** : 휴대 전화를 비롯한 휴대용 장치를 위한 운영 체제와 미들웨어, 사용자 인터페이스 그리고 표준 응용 프로그램을 포함하고 있는 소프트웨어 스택이자 모바일 운영 체제
- **API** : Application Programming Interface
- **APP, 앱** : Application
- **(Background) Service** : 안드로이드 애플리케이션의 핵심 구성요소 중의 하나로서 사용자에게 보이지 않는 영역에서 데이터를 처리하거나 기타 작업을 처리한다.
- **Bookmark** : 영상의 특정 시점에 대한 일종의 레이블(Label)이며, 이 레이블에는 해당 시점에 대한 링크가 포함되어 있어 영상의 특정 시점으로 곧바로 이동할 수 있게한다.

- **DB** : Database
- **GUI** : Graphic User Interface
- **Keyword Spotting** : 음파 형태의 입력 데이터를 분석하여 특정 단어가 발화되었는지를 판단하는 기술
- **OS** : Operating System
- **ScrollView, 스크롤뷰** : 안드로이드 애플리케이션의 화면 구조 중에 하나로서 해당 구조 내부에 존재하는 다른 화면 구조가 스크롤 될 수 있도록 한다.
- **Speech Recognition** : 사람이 말하는 음성 언어를 컴퓨터가 해석해 그 내용을 문자 데이터로 전환하는 기술
- **SRS** : Software Requirement Specification, 소프트웨어 요구사항 명세서
- **STT** : Speech-to-Text. Speech Recognition을 나타내는 또다른 표현
- **Thumbnail, 썸네일** : 사진의 축소판이며 사진을 탐색하면서 알아 보기 쉽게 만들어 주며 그림을 일반 문자열 색인과 같게 취급한다.
- **Toast Message, 토스트** : 안드로이드 기기에서 나타나는 작은 팝업으로 작업에 관한 간단한 피드백을 제공한다.
- **TTS** : Text-to-Speech. 화면의 텍스트를 음성 언어로 변환하여 말소리의 음파를 기계가 자동으로 만들어 내는 기술
- **UI** : User Interface
- **URL** : Uniform Resource Locator. 네트워크 상에서 자원이 어디 있는지를 알려주기 위한 규약
- **Wake-up word** : Keyword Spotting이 판단하는 특정 단어
- **Widget, 위젯** : 애플리케이션 사용자가 상호작용하는 인터페이스 요소
- **YouTube, 유튜브** : 미국의 동영상 공유 플랫폼. 전 세계 최대 규모의 동영상 공유 사이트로서, 사용자가 영상을 시청, 업로드, 공유할 수 있다.

1.6 개요 (Overview)

본 문서의 구성은 다음과 같다. 2장에서는 ‘쿠키야!’ 애플리케이션의 전반적인 개요를 알아본다. 이를 위해 프로젝트 결과물의 향후 전망과 결과물의 대략적인 기능, 사용자의 특성, 제약 조건, 가정과 의존성 등의 각 절로 나누어 살펴본다. 3장에서는 ‘쿠키야!’ 애플리케이션의 상세 요구사항을 알아본다. 외적 인터페이스 요구사항과 기능적 요구사항, 성능 요구사항, 설계 제약 조건, 소프트웨어 시스템 특성, 기능별 구현 우선순위를 각 절에서 살펴본다.

2. 종합 기술 (Overall Description)

2.1 프로젝트 결과물의 전망 (Product Perspective)

2.1.1 시스템 인터페이스 (System Interfaces)

시스템은 Android OS(Jelly Bean, API 16 이상)가 설치된 스마트 디바이스가 요구되며, 애플리케이션 상의 입력은 GUI 인터페이스를 위한 스마트 디바이스의 터치스크린과 음성 명령어 인식을 위한 기기 내부의 마이크로 구성된다. 애플리케이션 상의 출력은 스마트 기기의 스크린과 기기 내부의 스피커로 구성된다.

2.1.2 사용자 인터페이스 (User Interfaces)

사용자 인터페이스는 애플리케이션의 다양한 기능을 수행하는 위젯과 이를 포함하고 있는 여러개의 액티비티로 구성된다. 각 기능에 대응하는 위젯은 버튼, 비디오 프레임, 대화상자 등으로 나타나며 액티비티간의 이동은 애플리케이션의 위젯뿐만 아니라 스마트 기기에 내장되어 있는 '뒤로가기' 버튼 등으로 수행할 수 있다. 또한 음성을 통해서 애플리케이션이 제공하는 기능을 수행할 수 있다.

2.1.3 소프트웨어 인터페이스 (Software Interfaces)

- OS : Android 4.1 (Jelly Bean, API 16) 이상의 운영체제
- 개발환경 : Android Studio 3.6.2
- 개발언어 : Java

2.1.4 운영 (Operation)

- 라이선스 : 본 애플리케이션은 영구라이선스 정책으로 운영한다.
- 업데이트 및 배포 : 업데이트가 필요한 경우 구글 플레이 스토어를 통한 업데이트를 진행한다.

2.2 결과물의 기능 (Product Functions)

본 프로젝트의 결과물은 다음과 같은 기능을 수행한다.

- (기능 1) 사용자의 음성을 인식하여 명령어로 전환.
- (기능 2) YouTube 영상에 대한 링크를 내부에 저장한 뒤 앱 내부에서 영상 재생.
- (기능 3) 기능 2에서 저장한 영상의 링크들을 사용자가 원하는 기준에 따라 각각의 재생목록으로 분류.
- (기능 4) 기능 3에서 저장한 영상을 음성으로 선택재생.
- (기능 5) YouTube 영상의 재생을 음성으로 조절할 수 있는 인터페이스 제공.
- (기능 6) YouTube 영상의 특정 시점에 북마크를 생성하고, 해당 북마크 시점으로 이동할 수 있는 인터페이스 제공.
- (기능 7) 기능 6에서 생성한 북마크들을 저장 후 번호를 부여하여 번호를 통한 재생.

2.3 사용자의 특성 (User Characteristics)

어플리케이션의 사용자는 스마트 디바이스를 원활히 사용할 수 있으며, 발화에 문제가 없다고 가정한다. YouTube 동영상시청을 하는 사람 모두가 사용자가 될 수 있다.

2.4 제약 조건 (Constraints)

다음의 사항을 준수한다.

- 사용하는 기술은 이미 널리 성능이 입증된 것을 사용한다.
- 별도의 라이선스가 필요하거나 저작권료 지불이 필요한 기술 또는 소프트웨어의 사용을 피한다. (단, 본 시스템이 필수적으로 필요로 하는 유익한 기술 또는 소프트웨어의 경우는 제외한다.)
- 사용자에게 보다 친숙하고 편리한 방향으로 결정한다.
- 시스템의 추후 확장성과 가용성을 고려한다.
- 소스코드 작성 시 추후의 유지보수를 고려하여 충분한 주석을 추가한다.

2.5 가정과 의존성 (Assumptions and Dependencies)

본 문서의 모든 시스템은 안드로이드 운영체제를 기반으로 설계 및 구현한 것을 가정하여 작성되었다. 따라서 모든 내용은 안드로이드 운영체제를 기준으로 작성되었으며, 타 운영체제에서 동일하게 적용되지 않을 수 있다. 이후 타 운영체제로 버전을 고려 할 수 있다. 또한 본 애플리케이션은 YouTube 영상에 대한 링크를 저장하고, 해당 링크를 이용하여 사용자에게 YouTube 영상을 앱 내부에서 보여주기 때문에 사용자의 스마트기기가 네트워크에 연결되어있어야 한다.

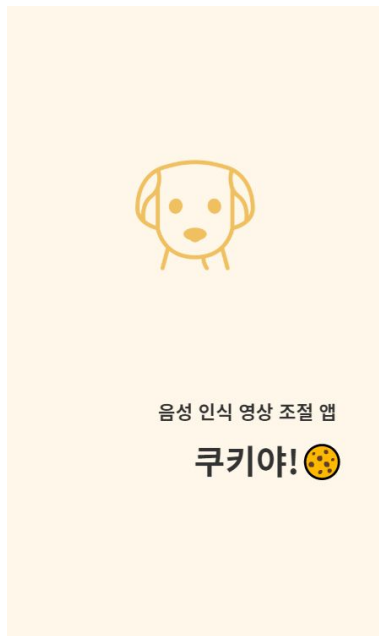
3. 상세 요구사항 (Specific Requirements)

3.1 외적 인터페이스 요구사항 (External Interface Requirements)

애플리케이션이 동작할 때 사용자에게 제공되는 User Interface를 설명한다. 여기서 보여지는 이미지는 설명을 위한 프로토타입들이며, 디자인과 세부사항은 추후에 변경될 수 있다. 설명의 편의를 위해서 화면을 나타내는 이미지 위에 빨간색 숫자를 표시하였다. 이 숫자는 설명을 위한 것이며 실제화면에서는 보이지 않는다.

3.1.1 로딩화면

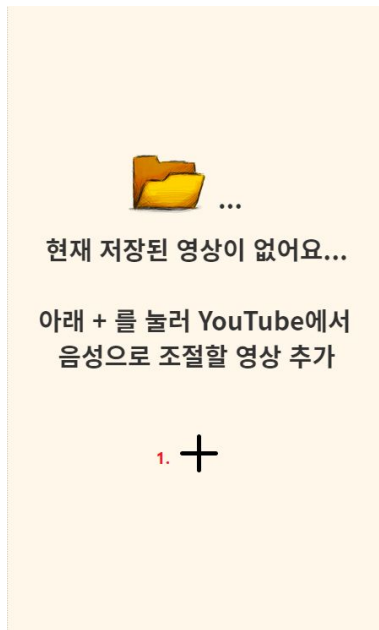
사용자의 안드로이드 기기에서 '쿠키야!' 앱을 실행시킨 경우 최초에 나타나는 화면이다. 이 화면은 다음과 같은 기능이 있다.



- 사용자에게 ‘쿠키야!’ 애플리케이션이 시작하는 중임을 나타내는 화면이다.
- 이 화면이 나타나는 동안 애플리케이션이 동작하기 위해서 필요한 작업을 내부적으로 수행할 수 있다.

3.1.2 영상목록 화면

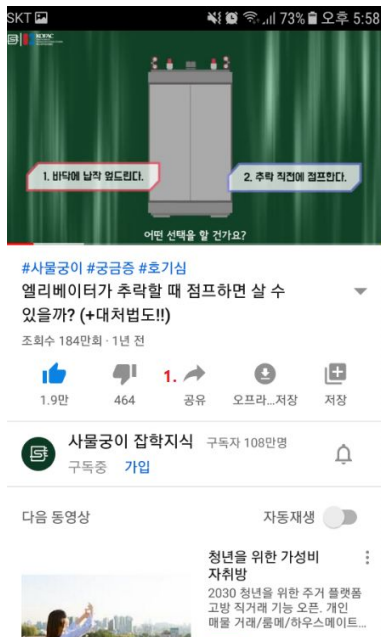
a. 애플리케이션에 저장된 영상이 없는 경우



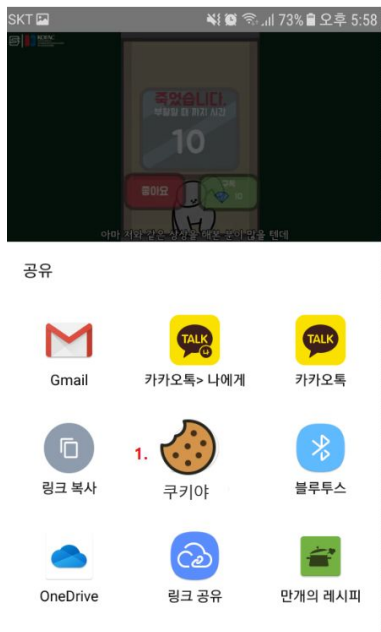
‘쿠키야!’ 애플리케이션을 설치한 후 처음 실행하거나, 애플리케이션에 저장한 YouTube 영상이 없는 경우에 보여지는 화면이다. 본 애플리케이션(‘쿠키야!’)은 YouTube 애플리케이션과 함께 동시 동작하는 애플리케이션이 아니다. 사용자는 YouTube에서 본인이 음성으로 조절하고 싶은 영상의 링크를 본 애플리케이션에 저장한 뒤, 본 애플리케이션이 독자적으로 동작하면서 영상을 조절하게 된다. 이 화면은 다음과 같은 기능이 있다.

- 숫자 1의 위치 오른쪽에 위치한 플러스(+) 기호 모양의 버튼이 존재한다. 이 버튼을 터치하는 경우 YouTube 애플리케이션이 실행된다. 만약 YouTube 애플리케이션이 존재하지 않는다면, 웹 브라우저가 실행된 뒤에 YouTube 웹사이트로 이동하게 된다.

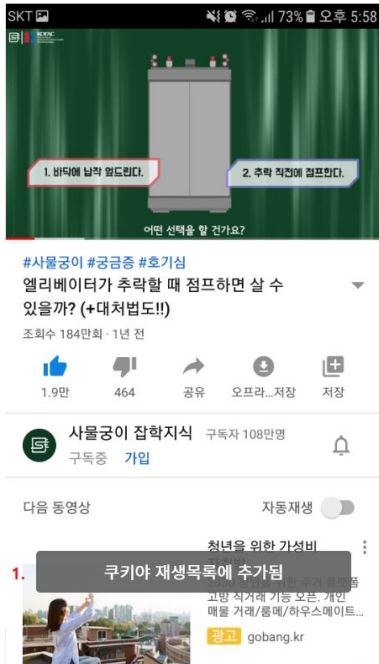
사용자는 YouTube 애플리케이션에 존재하는 임의의 영상을 다음과 같은 방법으로 본 애플리케이션('쿠키야!')에 저장할 수 있다.



- YouTube의 영상 재생화면에서 숫자 1의 오른쪽에 위치한 '공유' 버튼을 클릭한다.

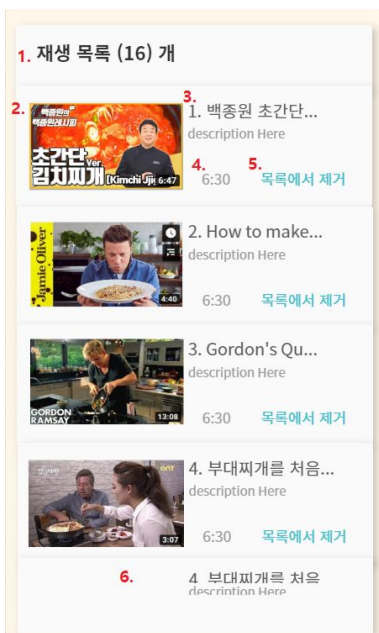


- 공유 기능에 대한 대화상자가 등장했을때, 숫자 1 위치 오른쪽에 나타난 것처럼 본 애플리케이션('쿠키야!')에 대한 아이콘이 나타난다. 해당 아이콘을 터치하면 '쿠키야!' 애플리케이션 내부에 현재 보고 있는 영상의 링크 (URL 형태의 Text)가 저장된다.



- 영상이 성공적으로 저장되면, 숫자 1 위치 오른쪽에 나타난 것처럼 안드로이드의 토스트 메시지를 통해서 사용자에게 영상이 성공적으로 저장되었음을 나타낸다.

b. 애플리케이션에 저장된 영상이 있는 경우

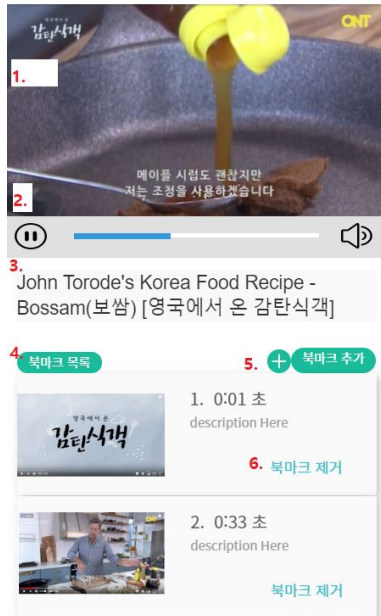


YouTube에서 음성으로 조절할 영상을 본 애플리케이션('쿠키야!') 저장한 후에는 왼쪽과 같은 화면이 나타난다. 이 화면은 다음과 같은 기능이 있다.

- 숫자 1 오른쪽에 나타난 것처럼 현재 애플리케이션에 저장된 영상의 링크 개수를 보여준다.
- 각 영상은 카드형태의 구조로 보여진다. 임의의 영상에 대한 카드를 나타내는 영역을 터치하는 경우 해당 영상을 재생하는 화면으로 이동한다.
- 숫자 2 오른쪽에 나타난 것처럼 영상의 썸네일이 카드 구조 왼쪽에 나타난다.
- 숫자 3 아래쪽에 나타난 것처럼 영상의 제목이 나타난다. 영상 제목 왼쪽에는 숫자가 부여되어 있는데, 해당 영상이 '쿠키야!' 애플리케이션에 저장된 순서대로 부여된 숫자이며, '이전 영상'과 '다음 영상'을 판단하는 기준이된다.
- 숫자 4 아래쪽에 나타난 것처럼 영상의 길이를 나타낸다.

- 숫자 5 아래쪽에 나타난 것처럼 '목록에서 제거' 버튼이 있어 사용자가 더이상 저장해두고 싶지 않는 영상을 해당 버튼을 눌러서 제거할 수 있다.
- 숫자 6이 나타나는 위치처럼 스마트기기의 화면에 모두 나타낼 수 없을 만큼 저장된 영상이 많다면 이 후 영상은 보여지지 않고, 사용자가 손가락을 위로 스윕핑할 경우 다음 영상에 대한 카드들이 보여지는 '스크롤뷰' 형태이다.

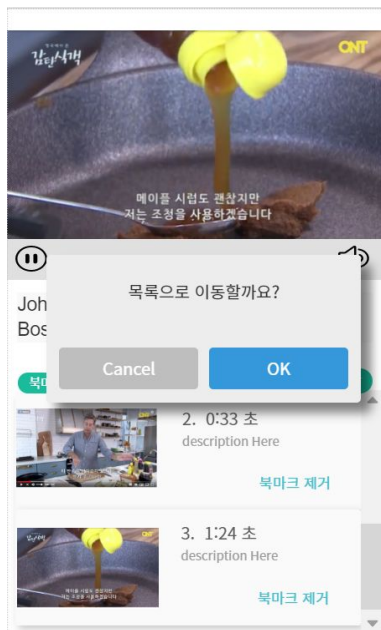
3.1.3 영상 재생화면



영상 목록화면에서 재생을 원하는 영상을 터치하면 나타나는 화면이다. 다음과 같은 기능이 있다.

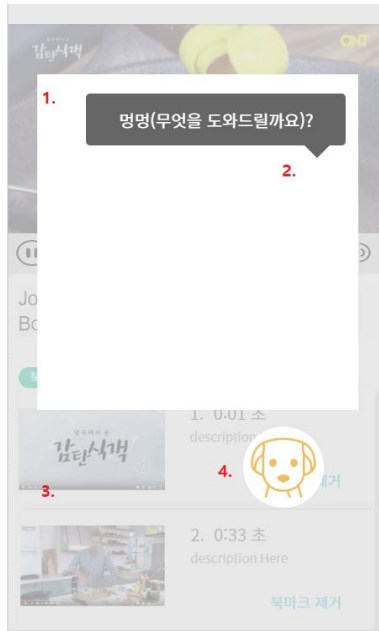
- 숫자 1 위치에 나타난 것처럼 비디오 플레이어를 이용해서 해당 영상을 플레이한다.
- 숫자 2 아래쪽에 나타난 것처럼 비디오 플레이어의 컨트롤러가 나타난다.
- 숫자 3 오른쪽에 나타난 것처럼 영상의 제목이 나타난다.
- 숫자 4 아래쪽에 나타난 것처럼 사용자가 지정한 북마크의 목록이 나타난다. 각 북마크는 카드형태로 나타나며, 썸네일, 북마크 시간의 위치, 북마크 제거 버튼이 나타난다. 여러 개의 북마크가 존재하여 기기 화면에 모두 나타나지 않는 경우 아래쪽에 위치한 북마크는 가려지게 되며, 위로 스윙했을 때 나머지 북마크가 나타나는 '스크롤뷰' 구조다.
- 숫자 5 오른쪽에 나타난 것처럼 '북마크 추가' 버튼이 있으며 해당 버튼을 터치하는 경우 플레이 되는 영상의 현재 시점에 대한 북마크가 생성된다.
- 숫자 6 오른쪽에 나타난 것처럼 '북마크 제거' 버튼이 있으며 해당 버튼을 터치하는 경우 북마크가 제거된다.

3.1.4 목록으로 이동하기 확인 대화상자



영상 재생화면에서 안드로이드 기기의 '뒤로가기' 버튼을 누른 경우 나타나는 확인 대화상자이다. OK 버튼을 누르는 경우 재생목록으로 이동하며, Cancel 버튼을 누르는 경우 원래 화면으로 복귀한다. 사용자가 영상을 보다가 실수로 뒤로가기 버튼을 누른 경우 의도하지 않게 목록 화면으로 넘어가지 않도록 하기 위해 대화상자가 나타난다.

3.1.5 Keyword Spotting이 일어난 화면



본 애플리케이션(‘쿠키야!’)을 실행하는 도중에 “쿠키야”라고 발화하면 애플리케이션의 백그라운드에서 실행되고 있던 Keyword Spotting 서비스는 Wake up word인 “쿠키야”를 인식한다. 그 다음 왼쪽과 같은 화면이 나타난다. 세부사항은 다음과 같다.

- 영상이 재생되고 있었다면, 영상을 일시정지한다.
- 숫자 1이 위치한 영역과 같이 대화상자가 나타난다.
- 숫자 2 위쪽에 나타난 것과 같이 사용자에게 어떤 명령을 내릴 것인지에 대한 프롬프트 메시지를 보여준다. 또한 메시지에 나타난 문장은 TTS(Text-To-Speech) 기능을 통해서 소리의 형태로 사용자에게 전달한다.
- 숫자 3이 위치한 영역과 같이 이전에 보고 있었던 화면은 흐려진다.
- 숫자 4 오른쪽에 나타난 것처럼 강아지 모양이 나타나 사용자에게 현재 음성인식 기능이 실행 중임을 나타낸다. 이 이미지는 음성인식 기능이 활성화된 경우에 모두 나타난다.

3.1.6 음성을 통해 영상 멈춤



Wake up word를 통해서 음성인식 기능을 활성화 시킨 뒤에 “영상 일시정지”와 같은 문장을 발화하는 경우 나타나는 화면이다. 다음과 같은 세부사항이 존재한다.

- 숫자 1 위쪽에 나타난 것처럼 사용자가 발화한 문장을 텍스트 형태로 보여줘서 사용자 본인이 발화한 내용과 일치하는지 확인할 수 있게 한다.
- 숫자 2 위쪽에 나타난 것처럼 유효한 명령어를 인식한 경우 “멍멍(알겠어요!)”와 같은 응답이 메시지 형태로 보여진다.
- 응답 메시지가 보여진 후에는 원래 화면으로 돌아가 영상을 일시정지 한다.

이 후부터 설명하는 외적 인터페이스에 대한 이미지는 3.1.6 절의 인터페이스 이미지와 구성이 동일하다. 문서의 공간을 절약하기 위해서 이후 절에서 설명하는 인터페이스의 이미지는 외적 인터페이스 설명이 끝나는 곳에 모아둔 뒤 캡션을 달아 어느 절에서 설명하는 인터페이스인지를 확인할 수 있도록 하였다.

3.1.7 음성을 통해 영상 다시 재생

Wake up word를 통해서 음성인식 기능을 활성화 시킨 뒤에 “영상 다시재생”과 같은 문장을 발화하는 경우 나타나는 화면이다. 화면 세부 사항은 3.1.6 절에서 설명한 것과 같으며, 응답메시지가 보여진 후에는 원래 화면으로 돌아가 영상을 다시 재생하게 된다.

3.1.8 음성을 통해 영상 간격 이동

Wake up word를 통해서 음성인식 기능을 활성화 시킨 뒤에 “10초 뒤로 이동해줘”와 같은 문장을 발화하는 경우 나타나는 화면이다. 화면 세부 사항은 3.1.6 절에서 설명한 것과 같으며, 응답메시지가 보여진 후에는 사용자가 지정한 간격만큼 영상의 플레이 시점을 옮긴다.

3.1.9 음성을 통해 북마크 생성

Wake up word를 통해서 음성인식 기능을 활성화 시킨 뒤에 “여기에 북마크 해줘”와 같은 문장을 발화하는 경우 나타나는 화면이다. 화면 세부 사항은 3.1.6 절에서 설명한 것과 같으며, 응답메시지가 보여진 후에는 현재 재생되고 있는 영상의 플레이 시점을 북마크로 기록한뒤, 북마크 목록에 추가한다.

3.1.10 음성을 통해 북마크로 이동

Wake up word를 통해서 음성인식 기능을 활성화 시킨 뒤에 “1번 북마크로 이동해줘”와 같은 문장을 발화하는 경우 나타나는 화면이다. 화면 세부 사항은 3.1.6 절에서 설명한 것과 같으며, 응답메시지가 보여진 후에는 사용자가 지정한 북마크가 나타내는 영상 플레이시점으로 이동한다.

3.1.11 음성을 통해 다음 영상 재생

Wake up word를 통해서 음성인식 기능을 활성화 시킨 뒤에 “다음 영상 재생해줘”와 같은 문장을 발화하는 경우 나타나는 화면이다. 화면 세부 사항은 3.1.6 절에서 설명한 것과 같으며, 응답메시지가 보여진 후에는 현재 재생되고 있는 영상의 다음 영상을 재생하기 시작한다.

3.1.12 음성을 통해 특정 영상 재생

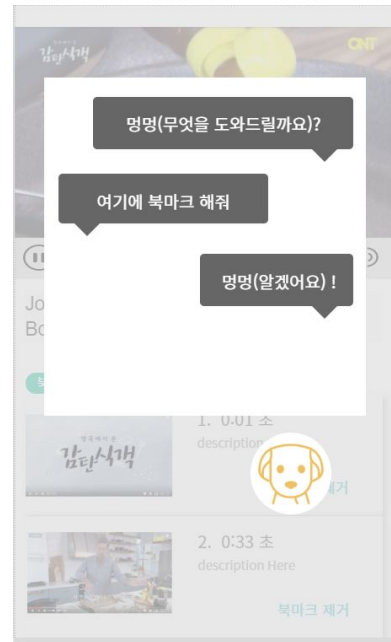
Wake up word를 통해서 음성인식 기능을 활성화 시킨 뒤에 “1번 영상 재생해줘”와 같은 문장을 발화하는 경우 나타나는 화면이다. 화면 세부 사항은 3.1.6 절에서 설명한 것과 같으며, 응답메시지가 보여진 후에는 사용자가 지정한 영상을 재생하기 시작한다.

3.1.13 음성을 통해 재생목록으로 이동

Wake up word를 통해서 음성인식 기능을 활성화 시킨 뒤에 “목록으로 이동해줘”와 같은 문장을 발화하는 경우 나타나는 화면이다. 화면 세부 사항은 3.1.6 절에서 설명한 것과 같으며, 응답메시지가 보여진 후에는 재생목록 화면으로 이동한다.

3.1.14 음성을 통해 프로그램 종료

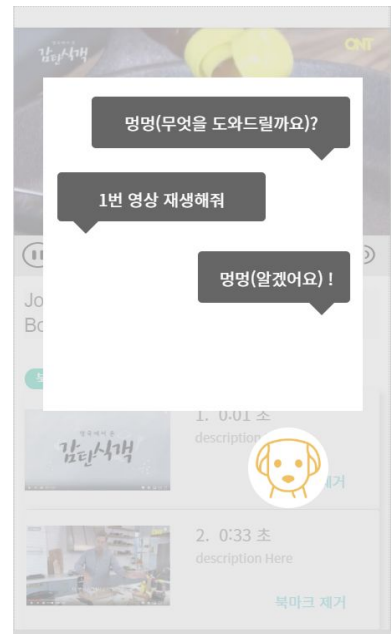
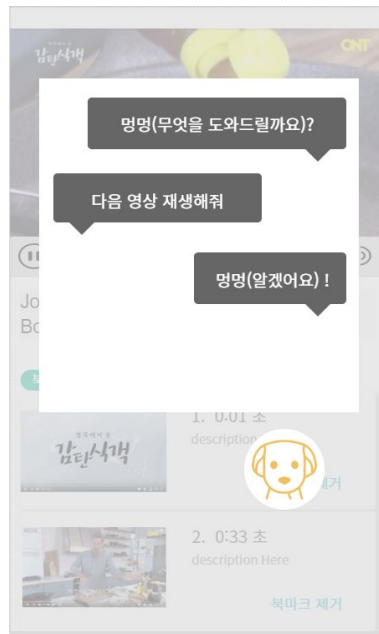
Wake up word를 통해서 음성인식 기능을 활성화 시킨 뒤에 “프로그램 종료”와 같은 문장을 발화하는 경우 나타나는 화면이다. 화면 세부 사항은 3.1.6 절에서 설명한 것과 같으며, 응답메시지가 보여진 후에는 애플리케이션을 종료한다.



3.1.7

3.1.8

3.1.9



3.1.10

3.1.11

3.1.12



3.1.13



3.1.14

3.2 기능적 요구사항 (Functional Requirements)

‘쿠키야! - 핸드프리 비디오 플레이어’ 애플리케이션은 이름 그대로 음성으로 유튜브 영상을 조작할 수 있는 어플리케이션이다. 사용자가 양손을 이용해 스마트 기기를 조작할 수 없는 경우 음성을 통해서 유튜브 영상을 조작한다. 사용자는 다음과 같은 기능을 사용할 수 있다.

3.2.1 특정 키워드 단어를 말하면, 해당 단어를 발화한 시점부터 음성인식 기능을 실행 - Keyword Spotting

스마트기기로 들어오는 모든 소리를 계속적으로 인식하여 동작한다면 많은 프로세싱 낭비가 발생할 수 있다. 따라서 비교적 간단한 키워드를 인식하는 프로세스를 로컬에서 실행하고 있다가 해당 키워드를 인식한 시점부터 음성인식 기능을 실행하도록 한다. (이처럼 특정 단어를 인식하는 알고리즘을 keyword spotting이라고 하며, 키워드는 wake up word라는 용어를 사용한다.) 아직 wake up word를 정하지 않았지만 문서 작성의 편의를 위해서 이후부터 “쿠키야” 라는 단어를 ‘쿠키야!’ 어플리케이션의 wake up word로 사용하겠다.

3.2.2 음성을 통해 유튜브 영상을 멈추고 재생

사용자가 “쿠키야, 영상 멈춰줘” 와 같은 문장을 발화하면 애플리케이션은 영상을 멈추고, “쿠키야, 영상 재생해줘” 와 같은 문장을 발화하면 애플리케이션은 영상을 다시 재생한다.

3.2.3 음성으로 영상 플레이 시점을 조절

사용자가 음성을 통해서 유튜브 영상의 플레이 시점을 조절(seek) 할 수 있는 기능이다. 아래에서 서술하는 종류의 기능을 제공한다.

a. 영상 플레이 시점을 특정 간격만큼 앞으로 이동 : 사용자가 “쿠키야, 10초 앞으로 이동” 과 같은 문장을 발화하면, 현재 보고 있는 영상의 플레이 시점을 10초 앞으로 이동시킨다

b. 영상 플레이 시점을 특정 간격만큼 뒤로 이동 : 사용자가 “쿠키야, 10초 뒤로 이동” 과 같은 문장을 발화하면, 현재 보고 있는 영상의 플레이 시점을 10초 뒤로 이동시킨다.

c. 영상 플레이 시점을 특정 시간대로 이동 : 사용자가 “쿠키야, 3분 10초 부터 재생” 과 같은 문장을 발화하면, 현재 보고 있는 영상의 플레이 시점을 사용자가 지정한 부분으로 이동시킨다.

위의 기능(a, b, c)을 명확히 하기 위해서 다음과 같은 세부사항을 기록한다.

- 기능 a, b, c의 경우 음성으로 영상 플레이 시점을 이동한 후에는 영상이 자동으로 재생된다. (설정을 통해서 영상 플레이 시점 조절 후에 영상이 일시정지 되도록 조절하는 기능을 고려할 수 있다.)
- 기능 a, b의 경우 사용자가 지정할 수 있는 간격의 종류로는 시간, 분, 초 단위로 한다.
- 기능 a, b의 경우 사용자가 여러가지 시간단위를 혼합하여 시간 간격을 조절할 수 있도록 한다. 예) “쿠키야, 1분 15초 앞으로 이동”

3.2.4 음성으로 영상에 북마크 설정 및 특정 북마크로 이동

사용자가 음성을 통해서 영상의 특정 시점에 북마크(Bookmark)를 설정할 수 있도록 한다. 여기서 북마크(Bookmark)란 영상의 특정 시점에 대한 일종의 레이블(Label)이다. 이 레이블은 숫자(1, 2, 3, ... 과 같은 자연수)로 표시되며, 사용자가 음성을 통해서 영상 플레이 시점을 조절할 수 있는 기준이 된다. 아래에서 서술하는 종류의 기능을 제공한다.

a. 영상이 재생되고 있는 현재의 플레이 시점에 북마크 설정 : 사용자가 “쿠키야, 여기에 북마크”와 같은 문장을 발화하면, 영상이 플레이 되고 있는 해당 시점에 북마크를 설정한다.

b. 영상 플레이 시점을 특정 북마크 시점으로 이동 : 사용자가 “쿠키야, 1번 북마크로 이동”과 같은 문장을 발화하면, 현재 보고 있는 영상의 플레이 시점을 1번 북마크가 설정된 시점으로 이동시킨다.

위의 기능(a, b)를 명확히 하기 위해서 다음과 같은 세부사항을 기록한다.

- 기능 a의 경우, 사용자가 지정한 북마크를 나타내는 숫자는 북마크가 생성된 순서대로 1번부터 1씩 증가하여 지정된다.
- 기능 b의 경우, 영상 플레이 시점을 특정 북마크가 나타내는 시점으로 이동시킨뒤에는 영상이 자동으로 재생된다. (설정을 통해서 영상 플레이 시점 조절 후에 영상이 일시정지 되도록 조절하는 기능을 고려할 수 있다.)

3.2.5 사용자가 시청할 유튜브 영상을 애플리케이션 내부로 저장

본 애플리케이션(‘쿠키야!’)은 구글사에서 제작한 애플리케이션 ‘YouTube’와는 별개로 동작하는 독립적인 애플리케이션이다. 사용자는 유튜브 영상을 시청하다가 음성을 통해 컨트롤 하고 싶은 영상에 대해서 ‘YouTube’ 애플리케이션의 ‘공유하기’ 버튼을 누른 뒤, 공유 가능한 애플리케이션 목록에서 본 애플리케이션(‘쿠키야!’)을 선택한 후에 해당 영상을 본 애플리케이션에 저장해둔다. 이는 유튜브 영상에 대한 링크를 저장하는 것으로 영상의 원본 파일을 다운로드 하는 것이 아니다. 아래에서 서술하는 기능을 제공한다.

a. ‘YouTube’의 공유하기 버튼을 통해서 복사된 유튜브 영상의 링크를 애플리케이션 내부에 저장하는 기능

b. 저장된 유튜브 영상의 목록을 화면에 보여주는 기능

위의 기능(a, b)을 명확히 하기 위해서 다음과 같은 세부사항을 기록한다.

- 기능 a의 경우, 사용자가 원하는 기준에 따라서 영상을 분류할 수 있도록 폴더 기능을 구현한다.
- 기능 b의 경우, 저장된 영상에 대하여 유튜브 영상의 제목, 썸네일, 영상의 길이 정보를 보여주도록 한다. 또한 영상이 저장된 순서대로 1번부터 번호를 부여한 뒤, 영상에 부여된 번호 정보도 보여주도록 한다.

3.2.6. 음성을 통해서 ‘음성인식 유튜브 영상조작’ 애플리케이션을 조작

사용자가 음성을 통해서 본 애플리케이션 (‘쿠키야!’)을 조작할 수 있도록 한다. 아래에서 서술하는 기능을 제공한다.

a. 음성을 통해서 특정 순서의 영상을 재생 : 사용자가 “쿠키야, 1번 영상 재생”과 같은 문장을 발화하면 1번으로 지정된 영상을 재생한다.

b. 음성을 통해서 앞 또는 뒤의 영상을 재생 : 사용자가 “쿠키야, 다음 영상 재생”, “쿠키야, 이전 영상 재생”과 같은 문장을 발화하면 다음 또는 이전 영상을 재생한다.

c. 음성을 통해서 영상 목록 화면으로 이동 : 사용자가 “쿠키야, 영상 목록으로”와 같은 문장을 발화하면 해당 영상이 포함된 목록을 보여주는 화면으로 이동한다.

d. 음성을 통해서 애플리케이션 종료 : 사용자가 “쿠키야, 프로그램 종료”와 같은 문장을 발화하면 애플리케이션을 종료한다.

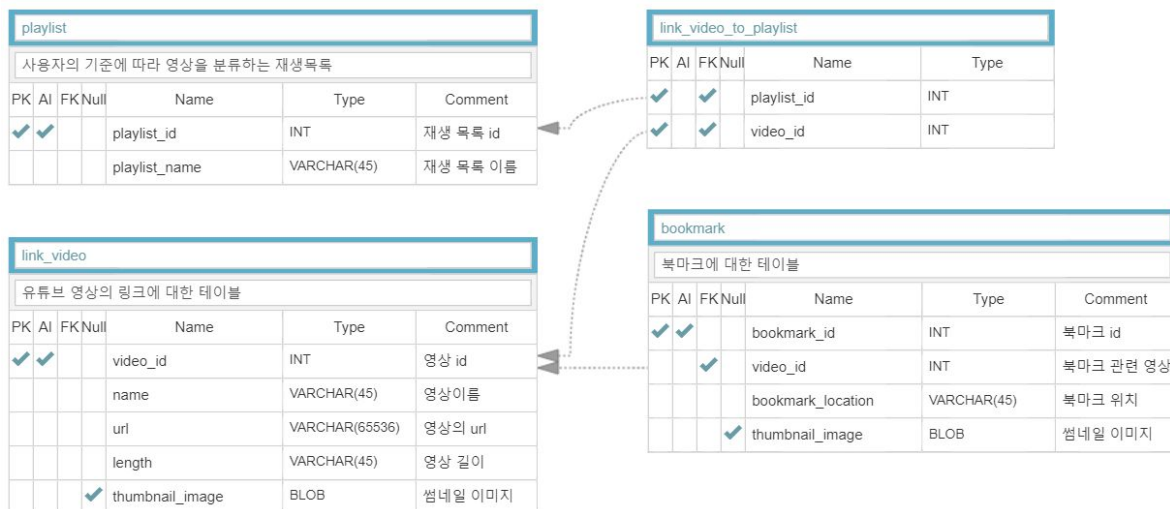
위의 기능(a, b, c, d)을 명확히 하기 위해서 다음과 같은 세부사항을 기록한다.

- 기능 a, b의 경우, 영상에 부여된 번호는 영상이 애플리케이션 내부에 저장된 순서 또는 영상이 포함된 폴더에 저장된 순서를 통해서 1번부터 차례대로 부여된다. 현재 영상에 부여된 숫자보다 숫자가 1 큰 영상이 ‘다음 영상’이 되고, 숫자가 1 작은 영상이 ‘이전 영상’이 된다.
- 기능 c의 경우, 폴더 구조를 구현한다면 현재 재생되고 있는 영상을 포함한 폴더의 목록 화면으로 이동한다.

3.3 성능 요구사항 (Performance Requirements)

본 시스템은 일반 사용자가 사용하는 스마트기기의 사양에서 안정적으로 구동되어야 한다. 또한 어떤 환경에서든 기기 외부에서 들려오는 사용자의 음성을 정확히 추출해낼 수 있어야 하며, 사용자가 음성을 통해 전달한 명령을 가급적 빠른 시간 내에 인식할 수 있어야 한다.

3.4 논리적 데이터베이스 요구사항 (Logical Database Requirements)



3.5 설계 제약사항 (Design Constraints)

본 장의 1절, 2절에서 언급한 내용을 바탕으로 시스템을 설계 및 구현한다. 그 외의 사항들은 개발자가 선호하는 방향을 선택하여 설계 및 구현하되, 단 다음의 사항을 준수한다.

- 사용되는 기술을 이미 널리 성능이 입증된 것을 사용한다.
- 별도의 라이선스가 필요하거나 저작권료 지불이 필요한 기술 또는 소프트웨어의 사용을 피한다(그러나, 애플리케이션 구현에 필수적인 기술 또는 소프트웨어의 사용은 예외로 한다).

3.6 소프트웨어 시스템 특성 (Software System Attributes)

- **신뢰도 (Reliability)** : 사용자가 터치 또는 음성으로 발화한 명령에 대해 정확한 기능을 수행
- **가용성 (Availability)** : 네트워크가 연결되어 있는 상황에서 사용자의 요청에 언제든지 허용 응답시간 내에 응답
- **유지가능성 (Maintainability)** : 시스템의 버그 보수 및 업데이트 용이성 극대화

3.7 기능별 구현 우선순위 (Functional Development Priority)

3.7.1 우선 구현사항

- (기능 1) 사용자의 음성을 인식하여 명령어로 전환.
- (기능 2) YouTube 영상에 대한 링크를 내부에 저장한 뒤 앱 내부에서 영상 재생.
- (기능 5) YouTube 영상의 재생을 음성으로 조절할 수 있는 인터페이스 제공.

3.7.2 선택 구현사항

- (기능 3) 기능 2에서 저장한 영상의 링크들을 사용자가 원하는 기준에 따라 각각의 재생목록으로 분류.
- (기능 4) 기능 3에서 저장한 영상을 음성으로 선택재생.
- (기능 6) YouTube 영상의 특정 시점에 북마크를 생성하고, 해당 북마크 시점으로 이동할 수 있는 인터페이스 제공.
- (기능 7) 기능6에서 생성한 북마크들을 저장 후 번호를 부여하여 번호를 통한 재생.