Text-Based Video Editing App – EasyUtube

4조
김동진 레쑤언휘 박세은 송민주 이은주 이종준

2019.09.09

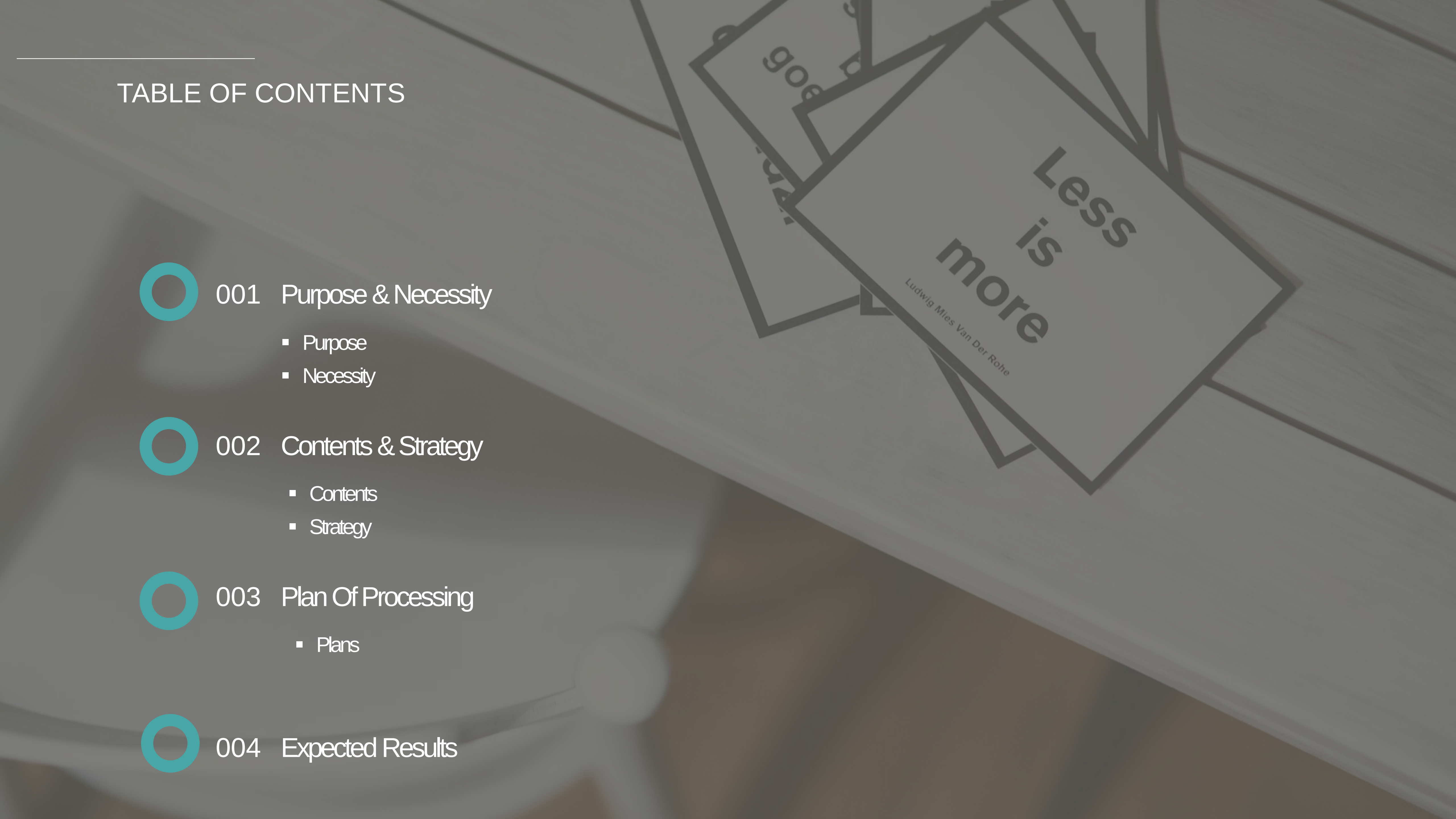


TABLE OF CONTENTS

001 Purpose & Necessity

- Purpose
- Necessity

002 Contents & Strategy

- Contents
- Strategy

003 Plan Of Processing

- Plans

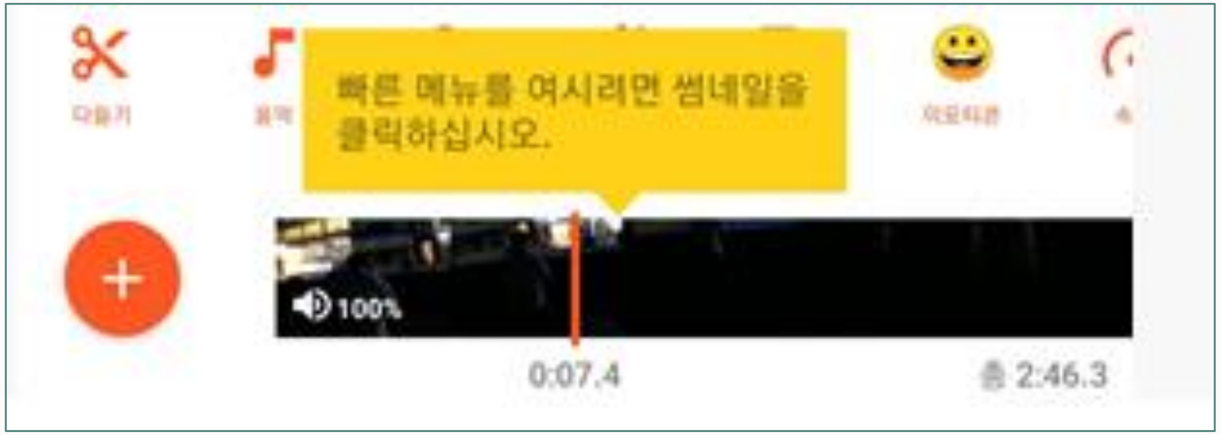
004 Expected Results

Part 1.

Purpose & Necessity

1.1 Purpose & Necessity

- **Disadvantages of the current video editing APP**

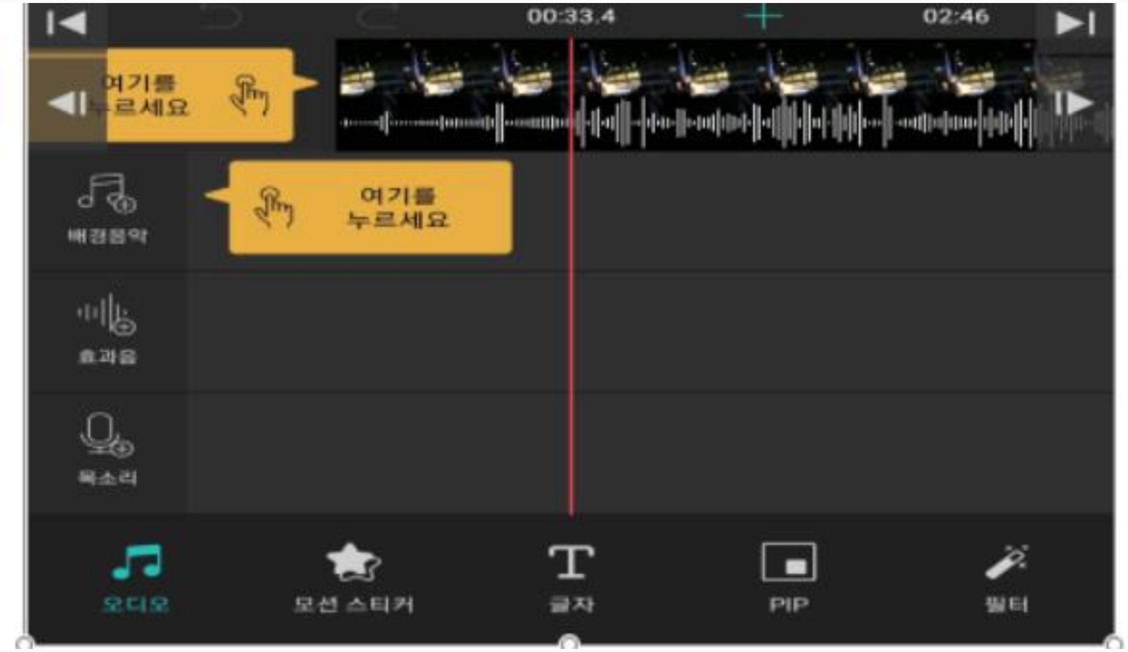


빠른 메뉴를 여시려면 썸네일을 클릭하십시오.

Ex) Viva Video(비바비디오)

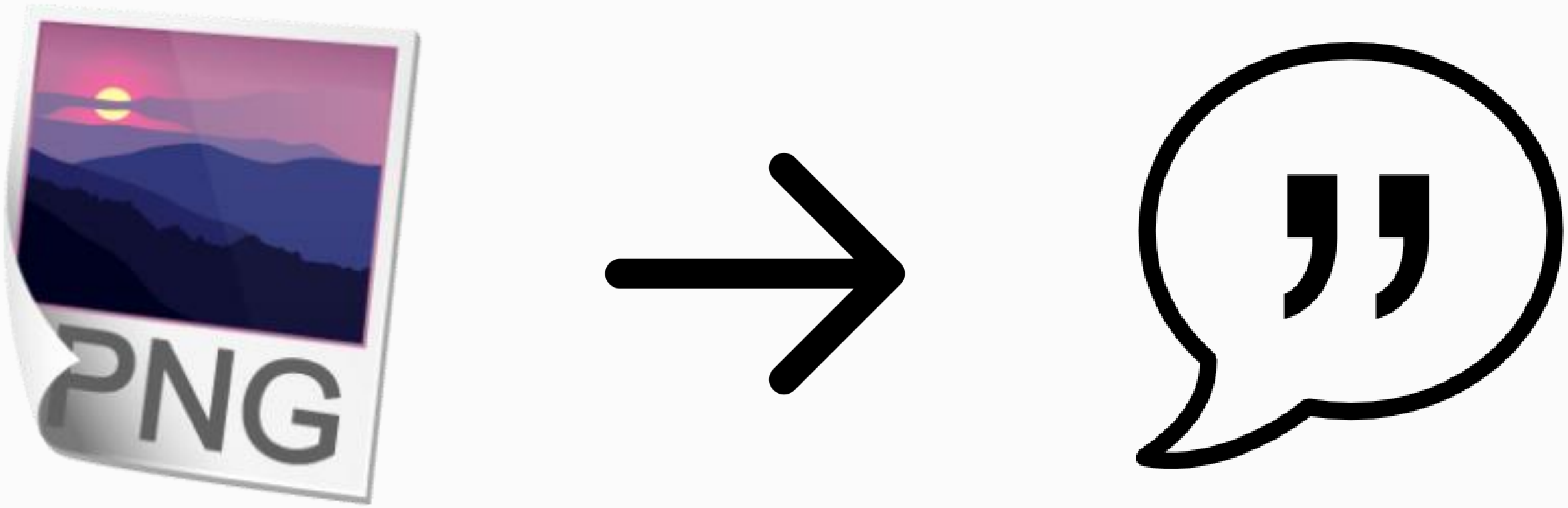
1.1 Purpose & Necessity

- **Disadvantages of the current video editing APP**



Ex) VLLO

1.1 Purpose & Necessity



Converting voice signals in images to TEXT based rather than image based

Part 2.

Contents & Strategy

2.1 Contents

- **Android UI/UX**

The UI/UX will be developed to fit the screen of the smartphone in the UI/UX of the existing clip/stick image

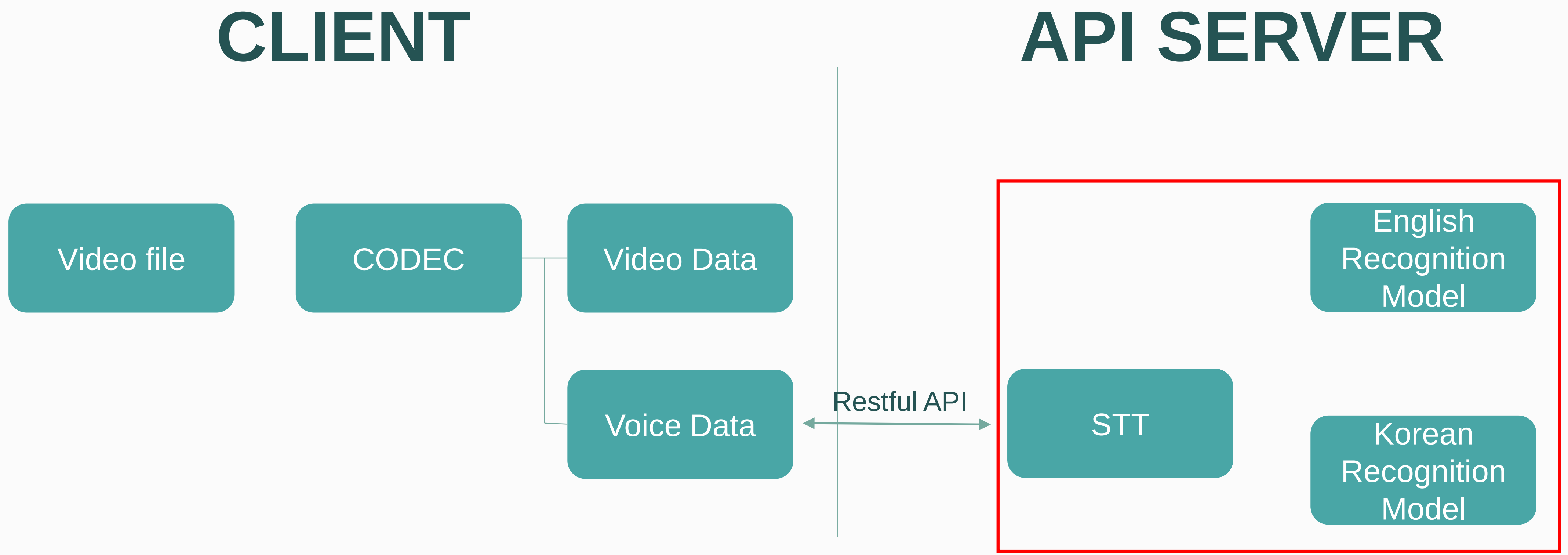
- **Video CODEC**

It is planning to develop video CODEC that works on Android OS

- **Voice Recognition Device**

Ability to convert English and Korean voice signals to TEXT and extract positional info

2.1 Contents



▲ System Configuration Diagram

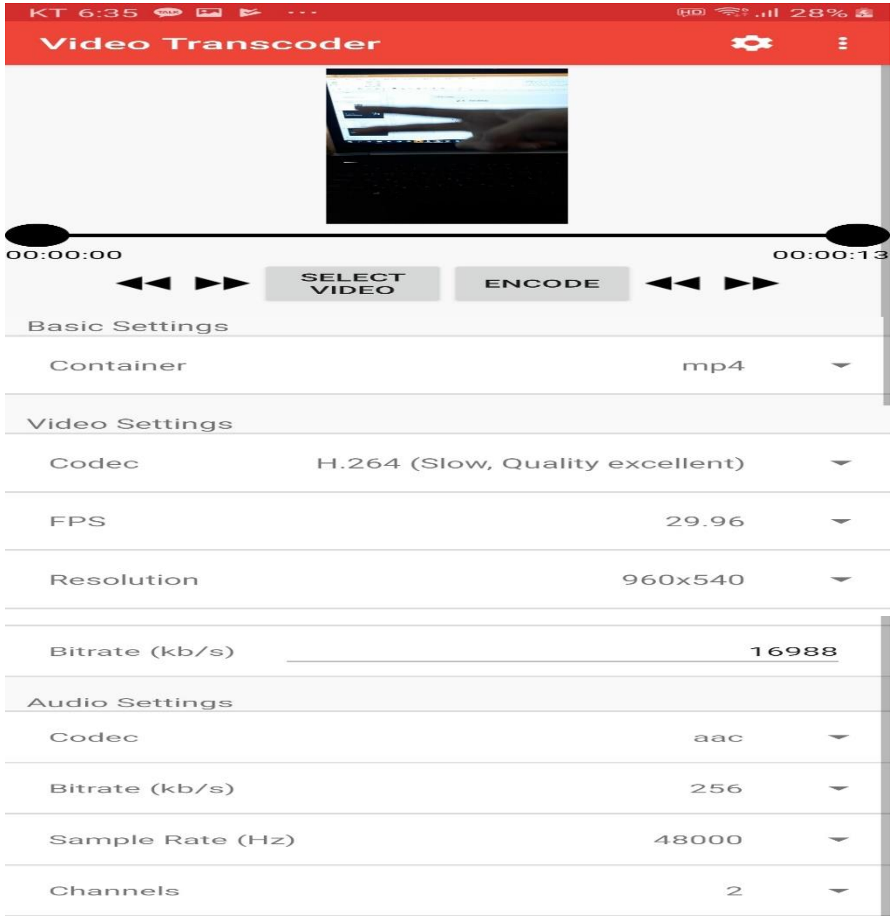
2.1 Strategy

(1) Video Application from Android apps

Play and Edit

2.1 Strategy

- Use Open Source



Basic Settings

Container mp4

Video Settings

Codec H.264 (Slow, Quality excellent)

FPS 29.96

Resolution 960x540

Bitrate (kb/s) 16988

Audio Settings

Codec aac

Bitrate (kb/s) 256

Sample Rate (Hz) 48000

Channels 2

2.1 Strategy

(2) FFMPEG application from Android App

- Audio and video separation
- Audio and video decoding in progress
- Play and Edit

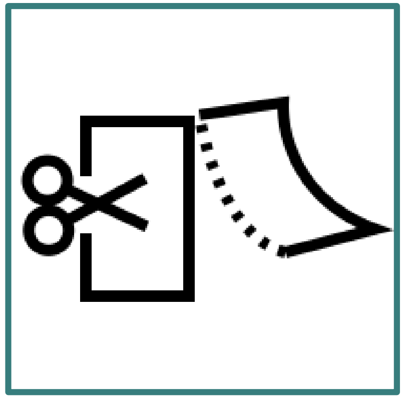
2.1 Strategy

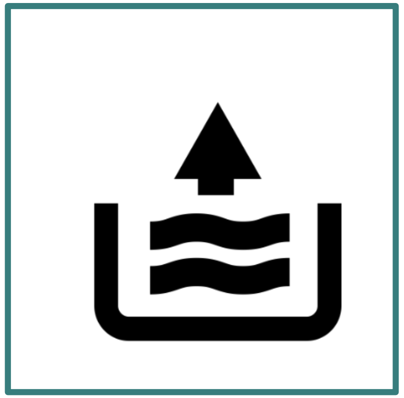
(2-1) FFMPEG

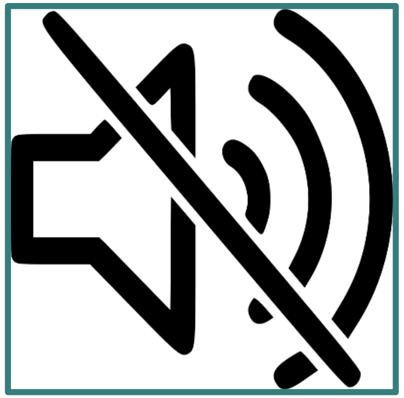
FFMPEG is a computer program that records and converts digital voice and video streams into various forms.

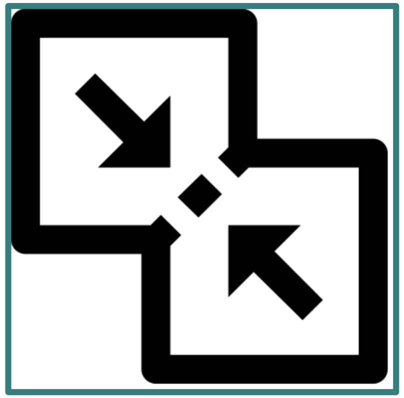
2.1 Strategy

(2-2) Example of FFMPEG







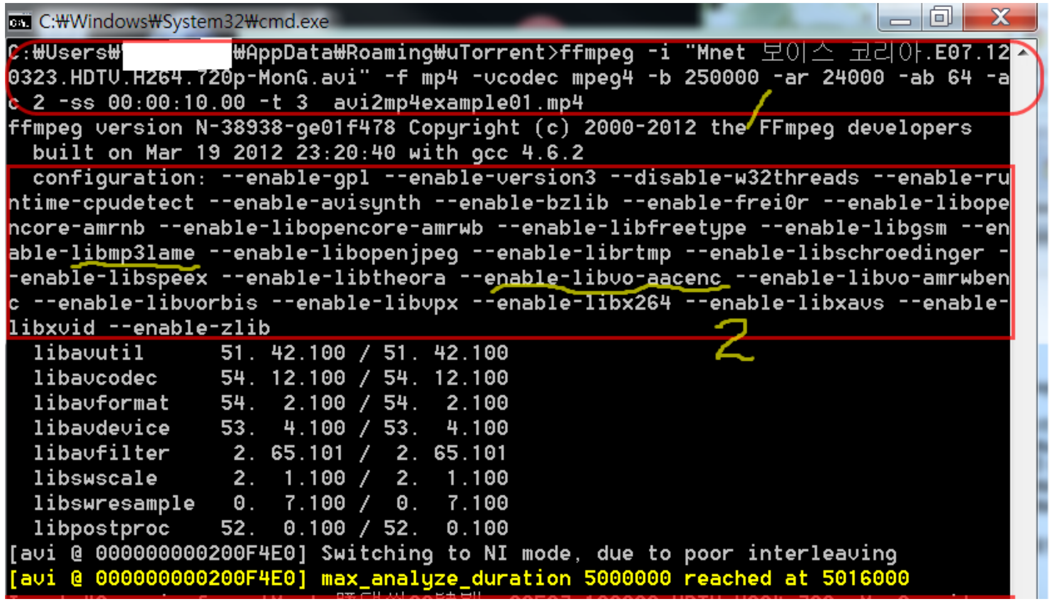


2.1 Strategy

(2-2) Example of FFMPEG

1) Import video for editing

2) Find the setting value of the cutting section



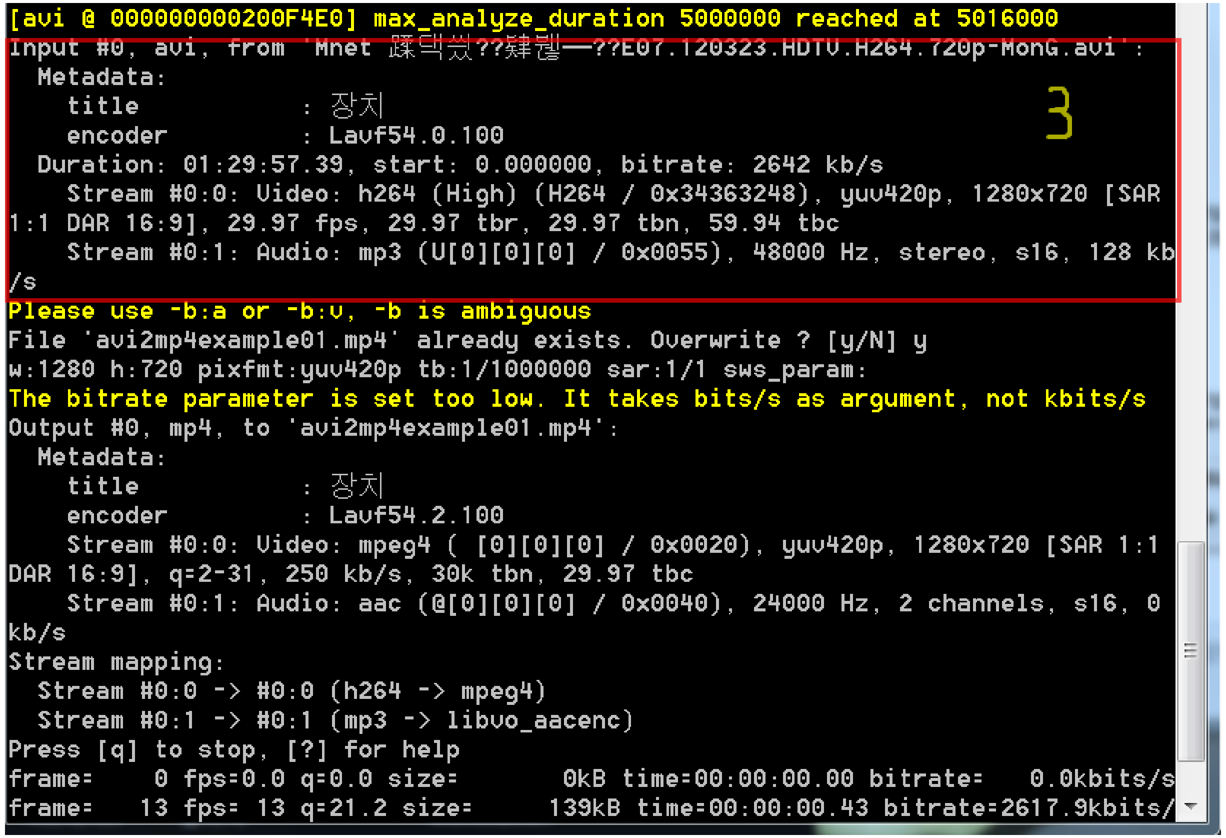
```
C:₩Users₩[redacted]₩AppData₩Roaming₩uTorrent>ffmpeg -i "Mnet 보이스 코리아.E07.12
0323.HDTV.H264.720p-MonG.avi" -f mp4 -vcodec mpeg4 -b 250000 -ar 24000 -ab 64 -a
c 2 -ss 00:00:10.00 -t 3  avi2mp4example01.mp4
ffmpeg version N-38938-ge01f478 Copyright (c) 2000-2012 the FFmpeg developers
  built on Mar 19 2012 23:20:40 with gcc 4.6.2
  configuration: --enable-gpl --enable-version3 --disable-w32threads --enable-ru
ntime-cpudetect --enable-avisynth --enable-bzlib --enable-frei0r --enable-libope
ncore-amrnb --enable-libopencore-amrwb --enable-libfreetype --enable-libgsm --en
able-libmp3lame --enable-libopenjpeg --enable-librtmp --enable-libschroedinger -
-enable-libspeex --enable-libtheora --enable-libvo-aacenc --enable-libvo-amrwben
c --enable-libvorbis --enable-libvpx --enable-libx264 --enable-libxavs --enable-
libxvid --enable-zlib
  libavutil      51. 42.100 / 51. 42.100
  libavcodec     54. 12.100 / 54. 12.100
  libavformat    54.  2.100 / 54.  2.100
  libavdevice    53.  4.100 / 53.  4.100
  libavfilter     2. 65.101 /  2. 65.101
  libswscale      2.  1.100 /  2.  1.100
  libswresample   0.  7.100 /  0.  7.100
  libpostproc    52.  0.100 / 52.  0.100
[avi @ 000000000200F4E0] Switching to NI mode, due to poor interleaving
[avi @ 000000000200F4E0] max_analyze_duration 5000000 reached at 5016000
```

출처 : https://windowx.tistory.com/entry/ffmpeg-encoding-%EC%98%88%EC%A0%9C

2.1 Strategy

(2-2) Example of FFMPEG

3) Cut the parts you want and create them to mp3 files



```
[avi @ 000000000200F4E0] max_analyze_duration 5000000 reached at 5016000
Input #0, avi, from 'Mnet 蹂댁씀??돩뒈—??E07.120323.HDTV.H264.720p-MonG.avi':
  Metadata:
    title           : 장치
    encoder         : Lavf54.0.100
  Duration: 01:29:57.39, start: 0.000000, bitrate: 2642 kb/s
    Stream #0:0: Video: h264 (High) (H264 / 0x34363248), yuv420p, 1280x720 [SAR
1:1 DAR 16:9], 29.97 fps, 29.97 tbr, 29.97 tbn, 59.94 tbc
    Stream #0:1: Audio: mp3 (U[0][0][0] / 0x0055), 48000 Hz, stereo, s16, 128 kb
/s
Please use -b:a or -b:v, -b is ambiguous
File 'avi2mp4example01.mp4' already exists. Overwrite ? [y/N] y
w:1280 h:720 pixfmt:yuv420p tb:1/1000000 sar:1/1 sws_param:
The bitrate parameter is set too low. It takes bits/s as argument, not kbits/s
Output #0, mp4, to 'avi2mp4example01.mp4':
  Metadata:
    title           : 장치
    encoder         : Lavf54.2.100
    Stream #0:0: Video: mpeg4 ( [0][0][0] / 0x0020), yuv420p, 1280x720 [SAR 1:1
DAR 16:9], q=2-31, 250 kb/s, 30k tbn, 29.97 tbc
    Stream #0:1: Audio: aac (@[0][0][0] / 0x0040), 24000 Hz, 2 channels, s16, 0
kb/s
Stream mapping:
  Stream #0:0 -> #0:0 (h264 -> mpeg4)
  Stream #0:1 -> #0:1 (mp3 -> libvo_aacenc)
Press [q] to stop, [?] for help
frame=    0 fps=0.0 q=0.0 size=       0kB time=00:00:00.00 bitrate=   0.0kbits/s
frame=   13 fps= 13 q=21.2 size=     139kB time=00:00:00.43 bitrate=2617.9kbits/
```

출처 : https://windowx.tistory.com/entry/ffmpeg-encoding-%EC%98%88%EC%A0%9C

2.1 Strategy

(3) Text Conversion with Deep Learning

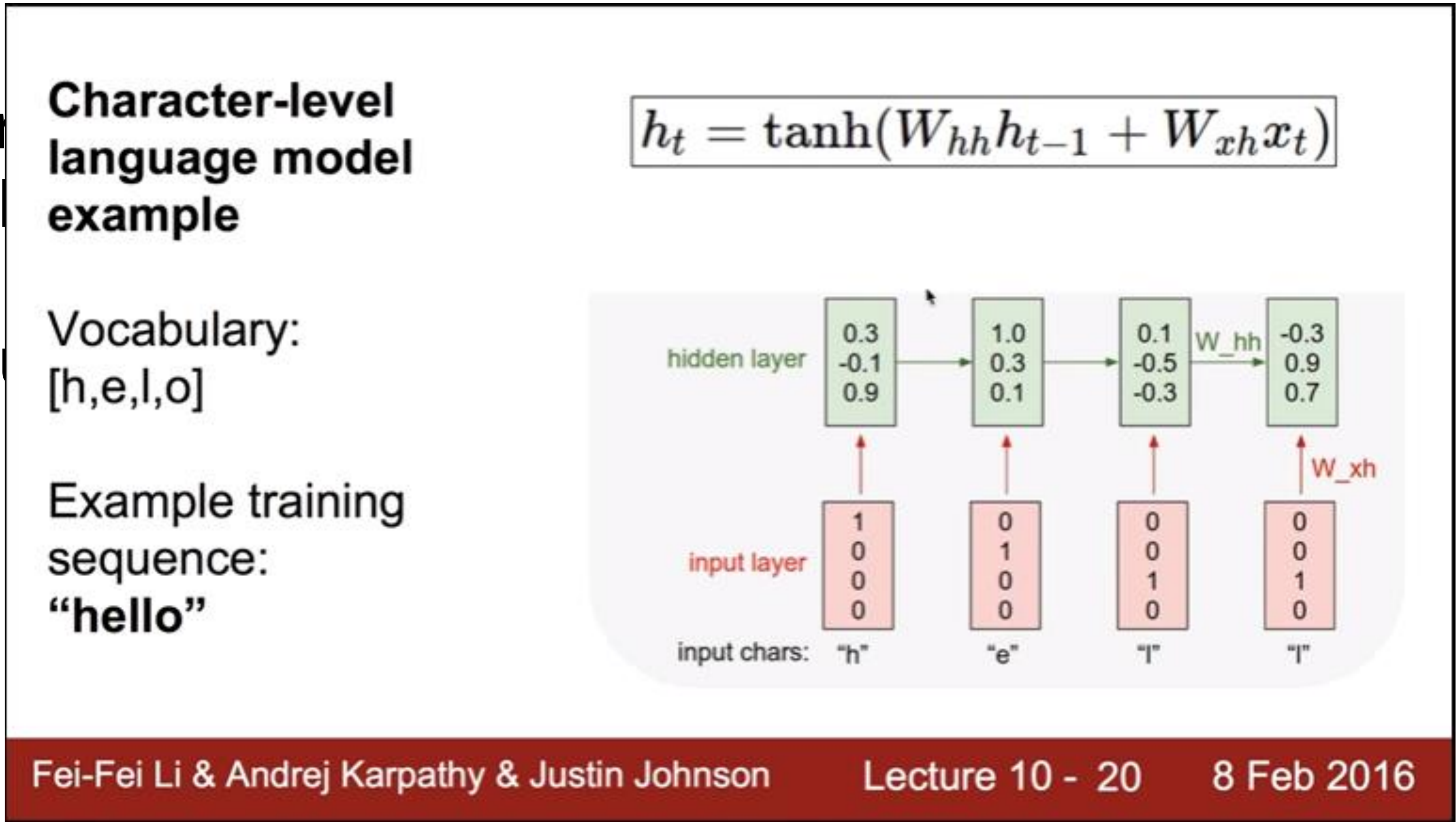
- Convert audio files to text using deep learning
- Use servers and clients
- Combine word-specific audio and video

2.1 Strategy

(3) Deep Learning Techniques for Text Conversion

<Recurr

- model
- voice e case of nat



Character-level language model example

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t)$$

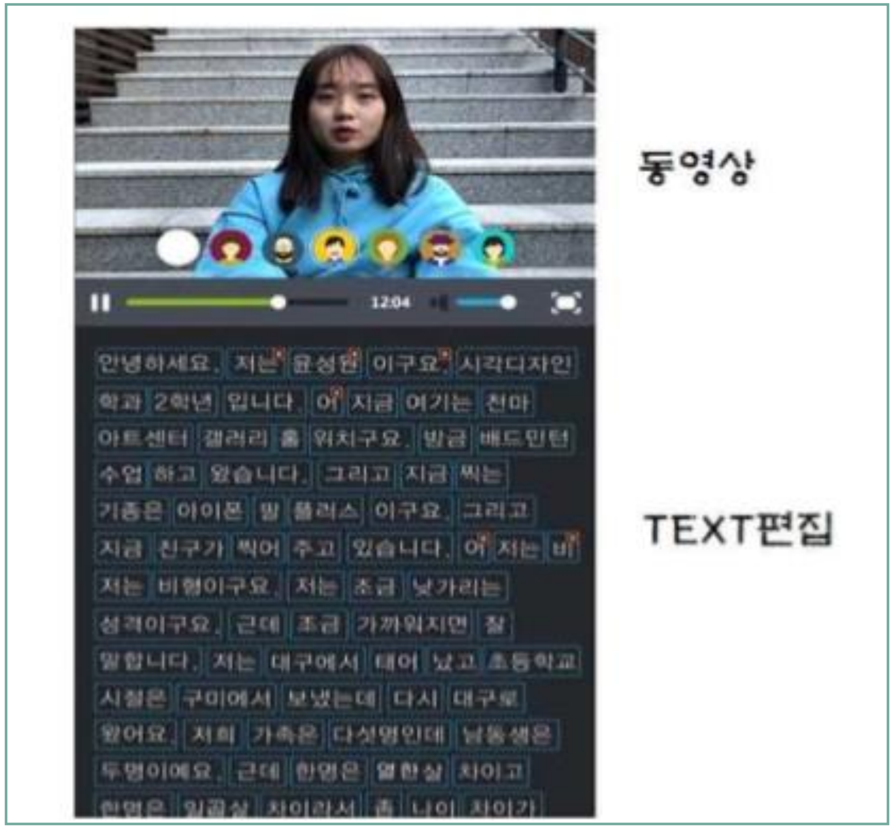
Vocabulary:
[h,e,l,o]

Example training sequence:
“hello”

Fei-Fei Li & Andrej Karpathy & Justin Johnson Lecture 10 - 20 8 Feb 2016

2.1 Strategy

(4) Text-based video editing



안녕하세요. 저는 윤성원 이구요. 시각디자인
학과 2학년 입니다. 어 지금 여기는 천마
아트센터 갤러리 홀 위치구요. 방금 배드민턴
수업 하고 왔습니다. 그리고 지금 찍는
기종은 아이폰 팔 플러스 이구요. 그리고
지금 친구가 찍어 주고 있습니다. 어 저는 비
저는 비형이구요. 저는 조금 낯가리는
성격이구요. 근데 조금 가까워지면 잘
말합니다. 저는 대구에서 태어 났고 초등학교
시절은 구미에서 보냈는데 다시 대구로
왔어요. 저희 가족은 다섯명인데 남동생은
두명이예요. 근데 한명은 열한살 차이고
한명은 일곱살 차이라서 좀 나이 차이가

Part 3.

Plan Of Processing

3.1 Plan

Period	Content
September Week 1-2	Team member composition and topic selection overall understanding and investigation of the subject Role sharing and implementation schedule plan by team member
September Week 3-5	Use ffmpeg in Android app
October Week 1-2	Play and edit Android app videos
October Week 3-5	intermediate inspection and fix bug Survey Deep Learning
November Week 1-3	Convert text using deep learning
November Week 4	Combine video by text
December Week 1	Modify Project Bug
December Week 2	Project testing and finishing Preparing Project Results Presentation

3.1 Plan

09/06 Meeting with Mentor

3.1 Plan

09/15 Conference

Part 4.

Expected Results

4.1 Expected Results

Thank You

4조