UDC 004.042

TRANSCRIPTION OF AUDIO AND VIDEO FILES IN THE INTERESTS OF LEARNING

Maglumzhanov M.A., Xenofontov D.A., Aishev A.A.

Military Engineering Institute of Radio Electronics and Communications of the Republic of Kazakhstan, Almaty, Republic of Kazakhstan

marlan84@mail.ru, xenofontov-dm@mail.ru, aisev@gmail.com

337

This article analyzes some software converters of audio and video clips into text form and considers the optimal way to transcribe a video clip in the interests of learning.

In the modern world, new information about education of various kinds appears every minute in large volumes, mainly in an accessible video format, for example, the YouTube channel. For each discipline, constant updating and replenishment of educational materials are required, on which the quality of education depends. There is a problem of text selection (transcription) from videos.

The solution to this problem is the use of the built-in means of the sound card of the stereo mixer of the Windows system with the use of a virtual cable driver for direct recording of the sound of a video clip, together with the use of the Speechpad transcription application online service. Which allows free and acceptable high-quality text conversion from video and audio with the following saving in text format.

Keywords: transcription; video clip; subtitles; audio; information; conversion; program; video hosting; stereo mixer; microphone; sound; buffer; sound card; noise.

There are people who do not use voice messages and videos, they will be uncomfortable. Viewing the text diagonally and understanding if there is something useful and interesting in it is often faster than listening to five minutes of audio or watching a video. Many educational institutions work with texts, as students learn mainly from books.

With the constant publication of video and audio content, there are times when for some reason you cannot see or listen to information, but you can read it.

Teaching staff in higher education institutions need this procedure to supplement educational materials in disciplines, write books, articles, scientific papers, etc.

Transcribing audio and video files is a tedious and not the most pleasant thing. But we live in the age of technology, which means that there are many opportunities for our services to transcribe audio and video files.

Currently, there are a huge number of converters for converting audio and video messages to text format. The highest quality option for translating audio into text is provided by professional transcription services. Their only drawback is that they are paid. But their advantages are more in comparison with free methods-high-quality transcription, predictable result. This can be verified by following some of the examples below.

So far, there is only one professional service for transcribing audio and video in Russian — Zapisano. In addition to Russian, it supports English, French, Spanish, Italian, German and other languages. Decoding of texts is carried out by specialists.

Payments for transcription start from 140 tenge per minute and depend on the urgency and complexity of the execution of the material. Text transcription is free.

Voco-a domestic program for translating video into text works on Windows OS. It will become an assistant for users who do not want to solve audio files themselves. You can pronounce the text into a microphone, get a ready-made audio recording or an audio track from a video. The program works with all popular audio formats. The program also performs video translation into text. The decoding quality is very high. The only and main minus is that the program is paid.

The Express Scribe program for Windows allows you to manually transcribe audio and video. To do this, the program has buttons for playing notes, pause and playback speed. If you do not want to type the text yourself, you can turn on the recording, set the desired speed, set the volume and run any voice input service, for example, Google Docs.

The disadvantage of the program is that there is no version in Russian. In addition, it is intuitive and easy to use.

The most popular video hosting in the world is the YouTube channel. This channel contains a lot of video materials of the educational process in the engineering and technical direction, from simple to complex, which is a source of audio and visual knowledge. There is a need to extract text from different videos [1].

If the video does not have subtitles, the "watch video transcription"option will not be displayed on YouTube. Accordingly, the above services do not provide a link to download subtitle files. If the video has a long introductory part with music or the sound quality is low, subtitles are usually not created. Also, the author of the YouTube channel can delete or delete them.

2. Open the recorder management window and make the "Cable Output" device the default, as shown in figure 1.
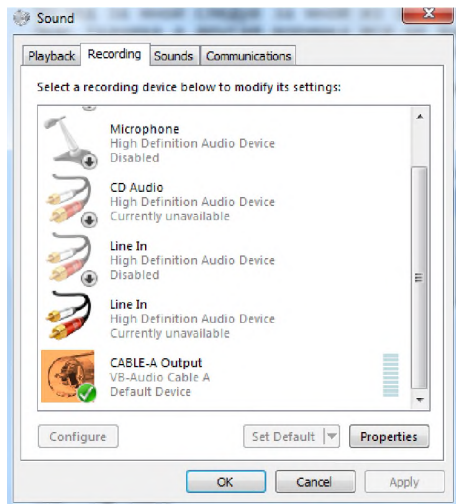
Figure 1 – Recorder control window

3. Open the playback device management window and make the "CABLE Input" device the default, as shown in figure 2.
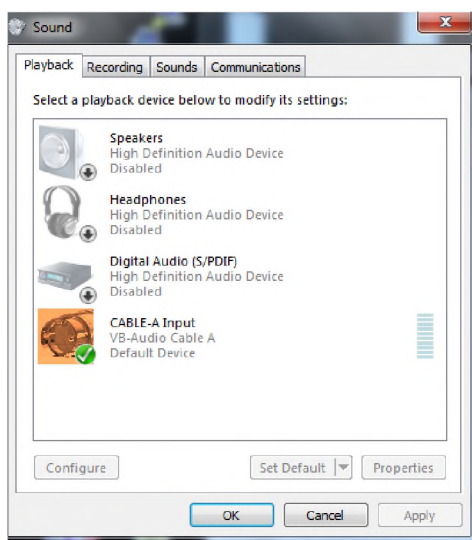

Figure 2 – Playback device management window

Now you can start transcribing. After these manipulations, the sound will go from the audio output to the recording, and the microphone will stop working. To return it to work, you need to return the changes you made (return everything).The audio-text translation panel opens after clicking the "transcription" button on the main page of the site, as shown in figure 3.
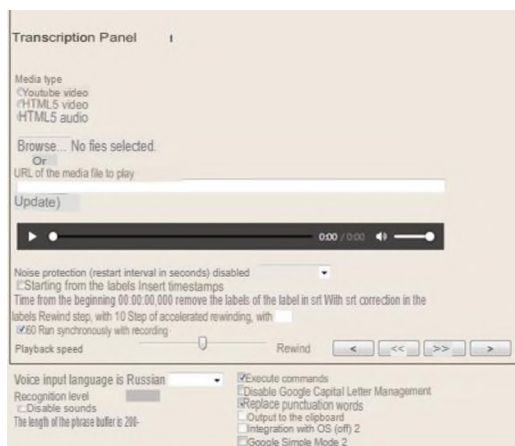

Figure 3 – transcription board

The "output to buffer and integration " checkboxes will be disabled. The" phrase buffer length " field prevents text from accumulating in the preview field (important when using virtual or physical cables).

The" noise protection" Switch is used to combat the freezing of the input in noisy videos. A good alternative to this switch is, for example, setting the "speech pause" switch in one second. The "pause speech" Switch can be seen in the user cabinet interface settings.

To download YouTube videos in the input field, you must enter its ID, as shown in figure 4.
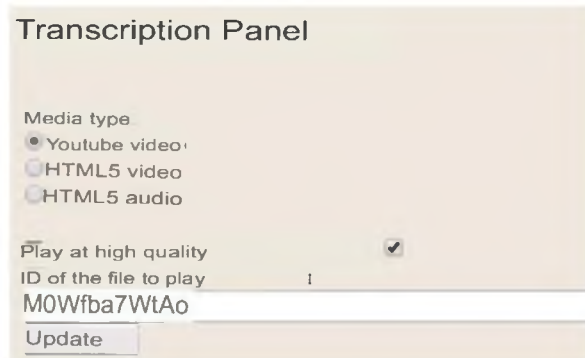

Transcription Panel

Media type
⦿ Youtube video
○ HTML5 video
○ HTML5 audio

Play at high quality   ✔
ID of the file to play   ₁
M0Wfba7WtAo
Update

Figure 4 – ID input field

Transcribing can be done synchronously, starting by checking the"run synchronously with recording" checkbox. When this checkbox is enabled, pressing the "Enable recording" button simultaneously turns on the player, and "stop recording" stops it [3].

In this case, the entire algorithm of work:

1) loading a roller or file into the player

2) ensuring sound from the roller to the microphone

3) setting the parameters, check the box "insert timestamps"

4) Click run.

There are errors in the text obtained by speech recognition. To correct it, the timestamps obtained during transcription are used. In this mode, you must disable the "run synchronously with a record" checkbox or use the start/stop Player key to expand notepad.

The correction algorithm is as follows:

1) the bulleted text first leads to a convenient view by sequentially pressing the "marks in SRT"and" marks in SRT " buttons

2) the" run from timestamps " checkbox is checked

3) the cursor is set to the desired position of the text

5) with the help of a hotkey or the" On/Off " button, the player listens to this part of the recording (the timestamp on the left side of the cursor is activated)

6) when using manual or notepad buttons, the text is edited

7) the marks are removed by clicking the "remove timestamps" button or the text is converted to SRT format (if credits are created for YouTube)

The disadvantage of this conversion method:

– constantly connected. internet resource.

– checking the correctness of word transformations from video and audio message to text after transcription.

Thus, the proposed method solves the problem of selecting text from audio and video. This method is necessary not only for teachers, but also for students (cadets), it can be used when there is a personal computer or laptop connected to the internet at home or at work.

It is very easy to add an installation algorithm. The necessary tools for this method are available on every computer and laptop, and even on smartphones. Speechpad online applications and virtual audio cable driver are freely distributed on the internet resource and are free. The choice of the transcription method will depend on the material capabilities, the urgency of the task and the sound quality.

And in some cases, if the recording quality is very poor and extraneous noise is heard, you will have to solve the transcription yourself or resort to the services of freelancers.

Reference
1. G. Shildt The complete guide to C# / G. Shildt Moscow: Williams, 2004 – 752 p.
2. Davies , K.H., Biddulph, R. and Balashek, S. (2002) *Automatic Speech Recognition of Spoken Digits*, J. Acoust. Soc. Am.24 (6) pp. 637— 642.
3. Лабор В.В. C#. Создание приложений для Windows – Минск: Харвест, 2003 – 385 с.

ТРАНСКРИБАЦИЯ АУДИО И ВИДЕОФАЙЛОВ В ИНТЕРЕСАХ ОБУЧЕНИЯ

Маглумжанов М.А., Ксенофонтов Д.А., Аишев А.А.

Военно-инженерный институт радиоэлектроники и связи Республики Казахстан, г. Алматы, Республика Казахстан

marlan84@mail.ru, xenofontov-dm@mail.ru, aisev@gmail.com

В данной статье проведен анализ некоторых программных преобразователей аудио- и видеороликов в текстовую форму и рассмотрен оптимальный способ транскрибации видеоролика в интересах обучения.

В современном мире ежеминутно в больших объемах появляется новая информация различного характера в сфере образовании, в основном в доступном формате видеоролика, например канал YouTube. По каждой дисциплине требуются постоянное обновление и пополнение учебных материалов, от которых зависит качество образования. Возникает проблема выделения текста (транскрибация) из видеороликов.

Решением данной проблемы является использование встроенных средств звуковой карты стереомикшера системы Windows с применением драйвера виртуального кабеля для прямой записи звука видеоролика совместно с использованием онлайн-сервиса приложения транскрибации Speechpad, которое позволяет бесплатно и допустимо качественно преобразовать текст из видео и аудио со следующим сохранением в текстовом формате.

Ключевые слова: транскрибация; видеоролик; субтитры; аудио; информация; преобразование; программа; видеохостинг; стереомикшер; микрофон; звук; буфер; звуковая карта; шумы.