**Assignment Report**

**Presentation in English II**

PhD student: Phan Xuan Thien

Student ID: 20141712

* Selected paragraphs for analysis: paper 2 – section 1.
* Original paragraphs (in which sentences are labeled by sequencing number):

“ (1)With current technologies, people can have many large broadcast video archives from TV programs. (2)This enables extracting more information or using archives more usefully than ever, by directly using high-dimensional information like video or audio data. (3)To make full use of video archives or to enable content-based video access to the archives, people need efficient methods to structuring or analyzing them.

(4)In video archives, there is some video segment that appears repeatedly, and those video segments can be used for analyzing the video archives [1]. (5)As a basic component of content-based video access, researchers tend to use feature-based similarity search for images and videos. (6)However, several papers have recently been published which state that similarity search is getting noisier and useless as the image/video archives are getting larger, instead, searching the “identical” image/video is becoming useful [2,3].

(7)In this paper, we are especially focusing on structuring news video archives. (8)There are many news channels that broadcast news programs 24 hours a day, and there are also broadcast stations which broadcast several hours of news program a day. (9)Those broadcast news videos occasionally have repeatedly used video segments. (10)Examples of identical video segments detected from an actual news video archive are shown in Fig. 1. (11)Images like Fig.1 (a) or (b) appears almost everyday in a collection of a news program. (12)Other images in Fig. 1 are appeared repeatedly because they are used in well-noticed topics, which broadcasted repeatedly for several days. (13)One of the reasons of this repetion is some kind of reference videos are usually inserted in news videos, and creating or newly acquiring reference video usually costs high.

(14)We noticed that most parts of video archives that share the same video segment has some relations, and such segments tends to be classified depending on their distribution in the broadcast time scale. (15)In this paper we search identical video segments in a news video archive and we implement a news video browser that uses search results of identical video segments, and investigate how the method of identical video segment detection can be used for structuring or extracting valuable information from news video archive.”

* Analysis on each paragraph:
* Paragraph 1:
* Main sentence (that carrying the main idea of paragraph): sentence (3), explain the need of efficient method for structuring and analyzing the video archives to make full use of them and enable content-based video access to the achieves.
* Main phrase (that carrying the main idea in the main sentence): “people need efficient methods to structuring or analyzing them”.
* Sentence (1) is poorly described. It can be rewrote as: “Current technologies allow people to access a variety of large broadcast video archives from TV programs”.
* (2): poorly described, ‘…or using archive…’ -> make confused about subject of the sentence. Corrected as: “This enables extracting more information and makes the use of archives become more effective than ever by directly using high-dimensional information like video or audio data”.
* (3): ‘…to structuring…’ -> wrong grammar. Rewrote as: “In order to make full use of video archives or to enable content-based video access to the archives, we need efficient methods for structuring or analyzing them“.
* Paragraph 2:
* Main sentence: sentence (6) -> describe the disadvantage of similarity search method and the usefulness of searching the “identical” image/video.
* Main phrase: ‘searching the “identical” image/video is becoming useful’.
* (4): ‘…there is some video segment that appears repeatedly…’ -> not well described, improved as: ‘video segments may appears repeatedly’.
* (6): ‘…several papers have recently been published which state that…’ -> unnecessary phrase, could be deleted; ‘…is getting noisier and useless…’ -> not well described, could make confuse about the meaning, also wrong grammar (should not use present continuous tense); ‘is becoming useful’ -> not optimal phrase. The sentence can be improved as: “However, similarity search become inefficient since it is affected by noise as the image or video archives are getting larger, instead, searching which is based on “identical” image/video is more effective [2,3]”.
* Paragraph 3:
* Main sentence: (7), indicate the problem that the paper focuses on.
* Main phrase: ‘structuring news video archives’
* (7): can be rewrote as: “This paper focuses on structuring news video archives”.
* (8): unnecessary since it does not contain any useful information for the paragraph -> should be deleted.
* (9): poorly structured, wrong grammar (should not use present perfect tense) -> rewrote as: “We observe that the videos which are used in existing broadcast news programs repeat using video segments occasionally”.
* (13): “repetion” -> wrong word, should use ‘repeat’ instead; The sentence can be improved as: “The reasons of this repeat is because some kinds of reference videos are usually inserted in news videos, and creating or newly acquiring reference videos usually costs high.”
* Paragraph 4:
* Main sentence: sentence (15), describe the works/contribution of the paper.
* Main phrase: “investigate how the method of identical video segment detection can be used for structuring or extracting valuable information from news video archive”.
* (15): too long sentence, poorly described and structured, it does not show and emphasize clearly the main contribution of the paper. Sentence (15) should be restructured and rewrote into 2 sentences as follow:
* “In this paper we propose a new method for identical video segment detection in news video archives for structuring them or extracting valuable information from them”.
* “We experiment searching identical video segments in a specific news video archive and use our video browser implementation, which uses search results of identical video segments, to evaluate the efficiency of the identical video segment detection method for structuring news video achieves or extracting valuable information from them”.
* Based on the above analysis, the original paragraphs can be rewrote/corrected as following (tracking changes function of MS Word is used here to emphasize the correction/change that is made):

Current technologies allow people to access a variety of large broadcast video archives from TV programs. This enables extracting more information and makes the use of archives become more effective than ever by directly using high-dimensional information like video or audio data. In order to make full use of video archives or to enable content-based video access to the archives, we need efficient methods for structuring or analyzing them.

In video archives, video segments may appear repeatedly and these video segments can be used for analyzing the video archives [1]. As a basic component of content-based video access, researchers tend to use feature-based similarity search for images and videos. However, similarity search become inefficient since it is affected by noise as the image or video archives are getting larger, instead, searching which is based on “identical” image/video is more effective [2,3].

This paper focuses on structuring news video archives. We observe that the videos which are used in existing broadcast news programs repeat using video segments occasionally. Examples of identical video segments detected from an actual news video archive are shown in Fig. 1. The images in Fig.1 (a) or (b) appear almost everyday in a collection of a news program. Other images in Fig. 1 appear repeatedly because they are used in well-noticed topics, which broadcasted repeatedly for several days. The reason of this repeat is because some kinds of reference videos are usually inserted in news videos, and creating or newly acquiring reference videos usually costs high.

We notice that most parts of video archives that share the same video segment has some relations, and such segments tends to be classified depending on their distribution in the broadcast time scale. In this paper we propose a new method for identical video segment detection in news video archives for structuring them or extracting valuable information from them. We experiment searching identical video segments in a specific news video archive and use our video browser implementation, which uses search results of identical video segments, to evaluate the efficiency of the identical video segment detection method for structuring news video achieves or extracting valuable information from them.