

## SEGMENTASI DAN PENGENALAN TEKS ARTIFISIAL UNTUK PENGINDEKSAN VIDEO

Adhitya Bram<sup>1</sup>, Eddy Muntina Dharma<sup>2</sup>, Fazmah Arif Yulianto<sup>3</sup>

<sup>1</sup>Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

---

### Abstrak

Video merupakan media informasi yang pada saat ini umum digunakan. Permasalahan muncul pada saat suatu jenis informasi hendak diambil dari kumpulan file video, bagaimana cara mendapatkan informasi dalam video database tersebut berdasarkan kandungan informasinya (content).

Salah satu solusi dalam menemukan informasi dalam file video tersebut adalah dengan melakukan segmentasi dan pengenalan kemunculan teks artifisial. Teks artifisial tersebut dapat dimanfaatkan sebagai indeks dalam pencarian informasi yang dikehendaki.

Hasil dari perangkat segmentasi dan pengenalan teks tersebut berupa file video baru yang sudah terindeks, indeks berupa informasi gambar yang sedang ditampilkan dalam bentuk karakter ASCII yang akan digunakan untuk melakukan pencarian informasi. Pengujian menunjukan sistem yang berhasil dibangun dapat melakukan segmentasi seluruh kemunculan teks dalam video sebesar 86,30%, keberhasilan pengenalan seluruh teks dalam video sebesar 77,42%, serta kemungkinan keberhasilan pencarian informasi dalam video sebesar 89,23%.

Kata Kunci : Teks artifisial

---

### Abstract

Video is one of information media that commonly used. We found problem when we want to get some kind of specific information from video files, how we can get a specific information from video database by its contents.

One of solution to find spesific information from a video file is to segmenting and recognizing artificial text occurences. Artificial text can be used as an index for specific information that we want to find.

The final output of this segmentation and text recognition application is a new indexed video file, index taken from artificial text that occurred in video image as video content information and represented in ASCII character to simplicate information finding. As the conclusion, system capable to segmenting 86,30% of text blocks occurency in video files, performing 77,42% precision for text blocks recognition, and probability of information retrieval by index is giving result 89,23% for each text occured in video files.

Keywords : Artificial text.

---

# 1. Pendahuluan

## 1.1 Latar Belakang

Video merupakan media informasi yang pada saat ini umum digunakan dan dapat menyampaikan informasi yang lebih jelas dibandingkan informasi yang disampaikan dalam bentuk teks atau audio. Permasalahan muncul pada saat suatu jenis informasi hendak diambil dari kumpulan file video seperti dalam suatu video database, bagaimana cara mendapatkan informasi dalam video tersebut berdasarkan kandungan informasinya (*searching by content*).

Salah satu solusi dalam menemukan informasi dalam file video tersebut adalah dengan memanfaatkan kemunculan teks yang ada dalam video, dimana kemunculannya pada video biasanya merupakan keterangan informasi yang sedang ditampilkan. Teks tersebut dapat dimanfaatkan sebagai indeks dalam pencarian informasi yang dikehendaki. Untuk dapat melakukan segmentasi dan pengenalan teks yang muncul pada video, metode utama yang digunakan dalam tugas akhir ini adalah deteksi kemunculan teks, segmentasi, dan pengenalan teks.

Hasil dari perangkat segmentasi dan pengenalan teks tersebut berupa file video yang ditambah dengan indeks yang berupa teks dalam karakter ASCII yang memuat informasi tayangan video yang sedang ditampilkan dan dapat dimanfaatkan untuk melakukan pencarian informasi.

## 1.2 Perumusan Masalah

Permasalahan utama yang harus dipecahkan dalam sistem video adalah sulitnya pencarian informasi dalam video yang ditampilkan berdasarkan contentnya, terutama pada file video dengan content yang bermacam-macam dalam jumlah yang banyak. Permasalahan tersebut dipecahkan dalam tugas akhir ini dengan membuat suatu perangkat lunak yang mampu melakukan proses segmentasi dan pengenalan teks artifisial yang muncul pada suatu file video sehingga dapat dilakukan proses pengindeksan file video.

### 1.3 Tujuan

Tujuan pembuatan tugas akhir ini adalah membangun perangkat lunak yang mampu melakukan proses segmentasi dan pengenalan teks artifisial yang muncul pada file video. Untuk merealisasikannya, kemampuan perangkat segmentasi dan pengenalan teks tersebut difokuskan dalam point-point sebagai berikut:

1. Mampu melakukan deteksi dan segmentasi kemunculan teks-teks artifisial dalam dalam file video.
2. Mampu melakukan binarisasi dari segmen-segmen gambar yang diduga sebagai kemunculan teks, hasil binarisasi ini digunakan sebagai input dari modul pengenalan teks.
3. Mampu melakukan pengenalan teks dari segmen-segmen gambar yang telah dipastikan sebagai teks sehingga diperoleh teks baru dalam representasi karakter ASCII.
4. Mampu membuat hasil akhir berupa file video terindeks.
5. Memiliki user interface berupa perangkat lunak untuk melakukan proses pengindeksan file video asli dan pencarian informasi yang diinginkan pada file video yang sudah terindeks.

### 1.4 Batasan Masalah

Batasan-batasan masalah dalam pembuatan tugas akhir ini adalah :

1. Teks yang mampu dikenali adalah karakter yang tampil secara horizontal.
2. Inputan perangkat lunak ini adalah file video tak terkompresi (\*.avi).
3. Mesin OCR yang digunakan adalah produk yang dijual bebas di pasaran, karena itu kesalahan internal OCR diluar tanggungjawab sistem.

## 1.5 Metodologi Penyelesaian Masalah

Metode penyelesaian masalah yang digunakan dalam pembuatan tugas akhir ini adalah :

1. Studi literatur tentang *video content retrieval*, *video processing*, dan *image processing* melalui buku, jurnal ilmiah, dan mailing list.
2. Penelitian dan percobaan mengenai deteksi posisi kemunculan teks dan deteksi validitas teks menggunakan teori *edge detection*, *image statistical operation*, dan *frame averaging*.
3. Penelitian dan percobaan teori *automatic tresholding* untuk menemukan metode mana yang tepat digunakan dalam proses binarisasi gambar yang diduga memuat teks dari sebuah frame video.
4. Penelitian dan percobaan perangkat *optical character recognition* untuk melakukan proses konversi dari segmen gambar yang memuat teks dalam sebuah frame video menjadi bentuk teks dalam standar ASCII.
5. Penelitian mengenai representasi dan format file video untuk membuat file video yang terindeks.
6. Pengujian kemampuan perangkat lunak terhadap segmentasi kemunculan teks, tingkat keberhasilan pengenalan teks, dan ketersediaan informasi indeks.

## 5. Kesimpulan dan Saran

### 5.1 Kesimpulan

Dari hasil penelitian ini dapat disimpulkan bahwa sistem yang telah dibuat dapat melakukan hal-hal sebagai berikut:

1. Mampu melakukan deteksi dan segmentasi kemunculan teks-teks artifisial dalam file video.
2. Mampu melakukan binarisasi dari segmen-segmen gambar yang diduga sebagai kemunculan teks, hasil binarisasi ini digunakan sebagai input dari modul pengenalan teks.
3. Mampu melakukan pengenalan teks dari segmen-segmen gambar yang telah dipastikan sebagai teks sehingga diperoleh teks baru dalam representasi karakter ASCII.
4. Mampu membuat hasil akhir berupa file video terindeks.
5. Memiliki user interface berupa perangkat lunak untuk melakukan proses pengindeksan file video asli dan pencarian informasi yang diinginkan pada file video yang sudah terindeks.

### 5.2 Saran

*Video Content Retrieval* merupakan objek penelitian yang cukup menarik bagi para informatikawan. Keberadaan software OCR sangat membantu dalam pelaksanaan penelitian ini, namun melihat bahwa software OCR didesain untuk teks cetak, maka ada beberapa bagian dalam teks video yang tidak mampu dideteksi dengan tepat oleh perangkat OCR sehingga riset mengenai OCR perlu untuk lebih ditingkatkan.

Harapan penulis sebagai *future works* adalah mempercepat proses indexing hingga mencapai *real-time* sehingga pengindeks yang terintegrasi dengan *player* dapat terwujud, sehingga tiap file video yang pernah dimainkan dapat terindeks.

## Daftar Pustaka

- 1 A.K. Jain and B. Yu, **Automatic text location in images and video frames**, *Pattern Recognit.*, vol. 31, no. 12, pp. 2055–2076, 1998.
- 2 A.K. Jain, M.N Murty, P.J Flynn, **Data Clustering – A Review**, *ACM computing surveys*, vol 31, no.3, September 1999.
- 3 Brian S. Morse, **Tresholding**, *Birmingham Young University*1998-2000. page 3.
- 4 C.Wolf and J.M.Jolion, **Extraction of artificial text in multimedia documents**, *Pattern Anal Applic* 6:309-326, Springer-Verlag London, 2003.
- 5 Celine Thillou and Bernard Gosselin, **Color Binarization for Complex Camera based Images**, *TCTS Lab, Faculte Polytechnique de Mons, Avenue Copernic, 700 Mons, Belgium*.
- 6 H. Li, D. Doermann, and O. Kia, **Automatic text detection and tracking in digital video**, *IEEE Trans. Image Processing*, vol. 9, pp. 147–156, Jan. 2000.
- 7 Kardi Teknomo, **K-Mean Clustering**, [www.planetsourcecode.com](http://www.planetsourcecode.com).
- 8 Luong Chi Mai, **UNESCO Module : Introduction to Image Processing and Computer Vision**, *Department of Pattern Recognition and Knowledge Engineering Institute of Information Technology Hanoi, Vietnam*.
- 9 M. A. Smith and T. Kanade, **Video skimming for quick browsing based on audio and image characterization**, *Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep. CMU-CS-95-186, July 1995*.
- 10 M.Junker and R.Hoch, **On the evaluation of document analysis components by recall, precision, and accuracy**, *Proceedings of ICDAR 1999, (1999) 713–716*.
- 11 N. Otsu, **A Threshold Selection Method From Gray-Level Histograms**, *IEEE Transactions on Systems, Man, and Cybernetics, Vol. 9, No. 1, pp. 62-66, 1979*.
- 12 Rainer Lienhart, **Localizing and Segmenting Text in Images and Videos**, *IEEE Transactions on Circuits and Systems for Video Technology, Vol.12, April 2002*.
- 13 R. Lienhart and W. Effelsberg, **Automatic text segmentation and text recognition for video indexing**, *Multimedia Syst.*, vol. 8, pp. 69–81, Jan 2000.

- 14 Toshio Sato, Takeo Kanade, Ellen K. Hughes, Michael A. Smith, Shin'ichi Satoh, **Video OCR: Indexing Digital News Libraries by Recognition of Superimposed Captions**, *Multimedia Engineering Laboratory, Toshiba Corporation Yanagi-cho, Saiwai-ku, Kawasaki 210-8501, Japan*.
- 15 V. Wu, R. Manmatha, and E. M. Riseman, **Finding text in images**, in *Proc. 2nd ACM Int. Conf. Digital Libraries*, Philadelphia, PA, July 1997, pp. 23–26.
- 16 Y. Zhong, K. Karu, and A. K. Jain, **Locating text in complex color images**, *Pattern Recognit.*, vol. 28, pp. 1523–1535, Oct. 1995.

