

## IDENTIFIKASI SHOT DAN SCENE VIDEO (STUDI KASUS INTERNET VIDEO BROWSING) VIDEO SHOT AND SCENE IDENTIFICATION (CASE STUDY INTERNET VIDEO BROWSING)

Kusuma Noer Adi Purnomo<sup>1, -2</sup>

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

---

### Abstrak

Video direpresentasikan sebagai serangkaian frame yang berurutan dan dihubungkan oleh suatu interval waktu. Internet video browsing adalah pencarian video dari sumber-sumber yang berada pada jaringan internet. Dengan semakin banyaknya video yang berada pada jaringan internet menyebabkan pencarian video menjadi masalah. Penyebab utamanya adalah tidak disediakannya indeks untuk suatu video yang dapat mempermudah dalam melakukan pencarian. Indeks dapat diperoleh dengan mengambil shot dan scene pada video yang bersangkutan.

Shot adalah pengambilan gambar (perekaman) pada kamera dalam satu waktu tanpa berhenti.

Scene adalah koleksi dari shot yang memiliki kedekatan hubungan secara visual dan waktu.

Setiap scene dapat mengandung suatu cerita berkenaan dengan video yang bersangkutan. Dari shot dan scene dapat disusun suatu struktur seperti layaknya daftar isi sehingga mempermudah dalam pencarian video. Karena itu, diperlukan identifikasi untuk mengambil shot dan scene suatu video. Tugas akhir ini membahas dan menganalisa bagaimana algoritma twin-comparison serta intelligent unsupervised clustering dapat melakukan identifikasi shot dan scene pada video.

**Kata Kunci :** identifikasi shot, identifikasi scene, daftar isi video, twin-comparison, intelligent unsupervised clustering.

---

### Abstract

A video is presented by sequence of frames and connected by constant time interval. Internet video browsing is a searching from video sources located in the internet network. The vast quantity of video in internet makes a problem for its searching. The main problem is no available index for videos to help searching. This index is got from video shot and scene.

Shot is an recording from camera without stop. Scene is a collection of a visually and temporally related shot. A single scene contains story about its video. From this, we can make a video table of content in order to help its searching. Hence, shot and scene identification is needed. This final project studies and analyzes how twin-comparison and intelligent unsupervised clustering algorithm indentify video shot and scene.

**Keywords :** shot identification, scene identification, video table of content, twin-comparison, intelligent unsupervised clustering.

Telkom  
University

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

*Internet video browsing* adalah pencarian video dari sumber-sumber yang berada pada jaringan internet. Teknik ini berkembang sangat pesat seiring penambahan bandwidth dan penggunaan video dalam berbagai bidang. Banyaknya video yang tersimpan dalam media sangat membantu dalam mengatasi kebutuhan akan hiburan dan informasi. Akan tetapi usaha yang dibutuhkan untuk mencari video yang diinginkan tidaklah mudah jika video yang tersedia sangatlah banyak dan beragam. Untuk mengatasi hal ini, video-video tersebut diklasifikasikan berdasar kategori tertentu misal, jenis video, waktu, tempat pembuatan, dan sebagainya. Kesulitan ini bisa dikurangi dengan cara tersebut.

Video yang diinginkan dapat didownload sesuai dengan kategori yang telah disediakan. Akan tetapi kendala yang dihadapi adalah besarnya ukuran file video, sehingga membutuhkan waktu yang lama untuk mendownloadnya. Disamping itu image yang disediakan sebagai representasi suatu video dirasakan kurang memberikan informasi yang cukup. Tak jarang timbul kekecewaan karena terjadi kesalahan dalam mendownload file video. Oleh karena itu, dibutuhkan beberapa bagian video yang benar-benar bisa memberikan gambaran cerita dari video tersebut.

Pengambilan bagian dari video dapat dilakukan secara manual menggunakan tool *video editing*. Hal tersebut tidak efisien karena video yang tersedia sangatlah banyak. Karena itu dibutuhkan suatu mekanisme untuk mengambil bagian dari video secara otomatis. Bagian yang diambil tidak boleh asal dan harus bisa memberikan gambaran cerita tentang video yang bersangkutan. Untuk dapat mengetahui gambaran cerita dari suatu video dengan cepat dan tepat dapat dilakukan dengan memanfaatkan daftar isi video. Orang yang ingin mendownload video tersebut dapat melihat daftar isinya terlebih

dahulu. Dari situ dia dapat memutuskan apakah dia mau mendownload video ini atau tidak.

Bagian video tersebut dapat diperoleh dari *video shot* dan *scene*. Secara umum shot merupakan pengambilan gambar (perekaman) pada kamera dalam satu waktu tanpa berhenti., sedangkan scene merupakan kumpulan dari shot yang dapat membangun cerita. Penggunaan shot saja belumlah cukup karena shot kurang dapat membangun cerita dibandingkan scene. Identifikasi shot dan scene diperlukan untuk mendapatkan kedua elemen tersebut. Dengan identifikasi shot dan scene dapat dibentuk suatu daftar isi video yang diharapkan bisa digunakan untuk mengetahui gambaran cerita video yang bersangkutan. Daftar isi tersebut kemudian digunakan dalam *internet video browsing* sehingga akan mempermudah pada pencarian video.

## 1.2 Perumusan Masalah

Dibutuhkan suatu daftar isi dari video untuk mengetahui deskripsi singkat dan lengkap suatu video. Dari daftar isi dapat diketahui apa isi sebenarnya video. Pembentukan daftar isi video membutuhkan suatu algoritma identifikasi shot dan scene. Identifikasi shot akan mencari batas shot pada frame-frame video. Batas shot akan ditemukan jika perbedaan suatu frame melebihi threshold yang ditentukan. Tetapi shot hanya merupakan cuplikan video yang bersifat fisik dan kurang bisa memberikan deskripsi tentang video tersebut. Karena itu beberapa shot kemudian dikelompokkan menjadi satu scene jika masih memiliki hubungan yang sama secara visual maupun waktu. Inilah tugas identifikasi video scene, yaitu mengelompokkan beberapa shot yang memiliki hubungan sama sehingga dapat dihasilkan suatu informasi yang berarti tentang video yang bersangkutan. Kemudian, hasil pendeteksian tersebut direpresentasikan sebagai daftar isi video untuk digunakan pada internet video browsing.

---

### 1.3 Tujuan Penulisan

Tujuan yang ingin dicapai dalam penulisan tugas akhir ini adalah sebagai berikut:

1. Mempelajari dan mengimplementasikan algoritma *Twin-Comparison* dan *Intelligent Unsupervised Clustering* untuk mengidentifikasi video shot dan scene menjadi sebuah perangkat lunak untuk dimanfaatkan pada internet video browsing.
2. Melakukan analisa aplikasi perangkat lunak yaitu ketepatan hasil deteksi yang meliputi jumlah shot dan scene yang dihasilkan, jumlah scene dan shot yang salah dan benar, waktu pemrosesan, serta pengaruh jenis video (video dengan karakter cepat atau lambat).

### 1.4 Batasan Masalah

Permasalahan dalam tugas akhir ini, memiliki batasan sebagai berikut:

1. Hanya menangani video digital dengan format AVI, MPEG, dan Real Video, WMV.
2. Pendeteksian dilakukan hanya pada frame, dan tidak berdasarkan pada audio.
3. Video yang digunakan adalah video berwarna (bukan hitam putih).
4. Tidak membahas encoding dan decoding video secara keseluruhan.

### 1.5 Metodologi Pemecahan Masalah

Metodologi penyelesaian masalah dalam penyusunan tugas akhir ini adalah sebagai berikut :

1. Studi literatur, tahap menambah wawasan dari buku, artikel, terutama dari naskah-naskah yang tersedia di internet dan sumber-sumber lainnya yang berhubungan dengan identifikasi shot dan scene video.
  2. Perumusan masalah, tahap mengenali masalah yang dikaji sebagai hasil dari studi literatur. Yaitu mencari penyebab sulitnya mendapatkan video yang tepat yang sesuai dengan keinginan.
-

3. Pengumpulan data, tahap mencari dan mengumpulkan data yang menunjang tugas akhir ini. Yaitu dengan mencari berbagai video yang dapat mewakili beberapa jenis video yang ada.
4. Analisa kebutuhan, tahap mengidentifikasi domain masalah, kebutuhan perangkat lunak yaitu untuk mengidentifikasi shot dan scene video dengan inputan beberapa jenis format video sehingga akan dihasilkan suatu bentuk informasi tentang shot dan scene video.
5. Desain, tahap merancang perangkat lunak yang akan dibangun. Yaitu dengan mendefinisi berbagai teknik dan prinsip yang diperlukan untuk implementasi algoritma shot dan scene video.
6. Implementasi, tahap dimana hasil analisa dan desain diimplementasikan dalam suatu bahasa pemrograman.
7. Pengujian, mengujicobakan aplikasi perangkat lunak menggunakan dataset yang ada dengan memasukkan parameter-parameter pengujian. Dataset berupa file video yang dapat mewakili jenis-jenis video yang ada, dengan format yang telah ditentukan pada batasan masalah.
8. Analisa hasil. Yaitu analisa terhadap ketepatan hasil deteksi yang meliputi jumlah scene dan shot yang salah dan benar, waktu pemrosesan, jumlah shot dan scene yang dihasilkan, serta pengaruh jenis video (video dengan karakter cepat atau lambat).

### 1.6 Sistematika Penulisan

Tugas Akhir ini disusun dengan sistematika pembahasan sebagai berikut :

#### **BAB I      PENDAHULUAN**

Berisi latar belakang, perumusan masalah, batasan masalah, tujuan pembahasan, metodologi penyelesaian masalah dan sistematika penulisan.

#### **BAB II     LANDASAN TEORI**

Bab ini memuat penjelasan mengenai teori dasar video, histogram warna, konsep *twin-comparison*, serta *Intelligent unsupervised clustering*

---

**BAB III ANALISA DAN PERANCANGAN SISTEM**

Membahas tentang analisa kebutuhan dan perancangan awal sistem dengan metode RUP menggunakan bahasa pemodelan UML

**BAB IV IMPLEMENTASI DAN ANALISA HASIL**

Bab ini memuat implementasi aplikasi yang dibangun, melakukan uji keakuratan dan evaluasi terhadap analisa

**BAB V KESIMPULAN DAN SARAN**

Berisi tentang kesimpulan dari hasil penelitian tugas akhir ini dan saran-saran untuk pengembangan lebih lanjut.



## BAB V

### KESIMPULAN DAN SARAN

#### 5.1. Kesimpulan

Berdasarkan percobaan dan analisa yang telah dibahas dan dilaksanakan pada bab IV, maka kesimpulan sementara yang dapat diperoleh sebagai berikut :

- Algoritma Twin Comparison berhasil mengidentifikasi shot boundary yang disebabkan oleh camera break dan dissolve dengan cukup akurat yaitu accuracy antara 80% - 90%.
- Dari percobaan yang dilakukan, penggunaan bin histogram = 4096 lebih baik daripada bin histogram = 512. Tetapi dari segi waktu, bin histogram 512 lebih cepat daripada 4096.
- Untuk pemilihan threshold shot yang memberikan hasil paling baik adalah threshold cut = 5 dan threshold GT = 8.
- Algoritma Intelligent Unsupervised Clustering berhasil membangun struktur video yang terdiri dari shot, group dan scene. Algoritma tersebut paling efektif pada film dengan karakteristik pelan dan pencahayaan terang. Dan kurang cocok pada video dengan karakteristik cepat atau gelap.
- Untuk mengurangi kesalahan identifikasi, user dapat memilih shot/scene mana saja yang akan dimasukkan dalam struktur video.

#### 5.2. Saran

Berikut ini saran-saran yang perlu dipertimbangkan untuk pengembangan aplikasi lebih lanjut:

- Identifikasi shot di sini masih terbatas pada camera break dan dissolve.
- Algoritma ini kurang efektif dalam mengidentifikasi shot boundary yang disebabkan beberapa transisi lainnya, seperti panning, zooming, fade-in, fade-out dan visual efek. Diharapkan pada pengembangan lebih lanjut dapat mengidentifikasi transisi tersebut.

- Pergerakan objek yang cepat dapat menyebabkan salah identifikasi. Diharapkan pada pengembangan lebih lanjut dapat menangani hal ini.



## DAFTAR PUSTAKA

- [1] Rui, Y., Huang, T.S., Mehrotra, S. 1999. *Constructing Table-of-Content for Videos*. ACM Multimedia Systems Journal, Special Issue Multimedia Systems on Video Libraries.
- [2] Zhang, H.J., Kankanhalli, A., Smoliar, S.W., 1993. *Automatic Partitioning of Full-motion Video*. Multimedia System.
- [3] Jordi M., Gabriel F., 1999. *Video Shot Boundary Detection Based On Color Histogram*. SIGMOD.
- [4] Ng, C.W., Lyu, M.R. 2002. *ADVISE : Advanced Digital Video Information Segmentation Engine*. The 11<sup>th</sup> International World Wide Web Conference (WWW2002).
- [5] Chen, S.-C., Shyu, M.-L., Zhang, C., Kashyap, R.L. 2001. *Video Scene Change Detection Method Using Unsupervised Segmentation and Object Tracking*. IEEE International Conference on Multimedia and Expo (ICME01).
- [6] Kadir, Abdul. 2004. *Panduan Pemrograman Visual C++*. Yogyakarta: Penerbit Andi Yogyakarta.
- [7] Larman, Craig. 1998. *Applying UML and Pattern*. Prentice Hall.
- [8] Rasheed, Z., and Shah, Mubarak. 2003. *A Graph Theoretic Approach for Scene Detection in Produced Videos*. Multimedia Information Retrieval Workshop.
- [9] Yoshitaka, A., Miyake, M. 2001. *Scene Detection by Audio Visual Features*. IEEE International Conference on Multimedia and Expo (ICME01).
- [10] Patel, N.V., Sethi, I.K. *Compressed Video Processing for Cut Detection*. IEEE Proc. Vision, Image and Signal Processing, vol. 143. 1996.
- [11] Rasheed, Z., Shah, Mubarak. 2003. *Scene Detection In Hollywood Movies and TV Shows*. IEEE Computer Society Conference.
- [12] Sethi, I.K., Patel, N. 1995. *A Statistical Approach to Scene Change Detection*. SPIE Conference Proceedings on Storage and Retrieval for Image and Video Databases III.

[13] Ngai, C.H., Chan, P.W., Yau, E., Lyu, M.R. 2002. *XVIP: An XML-Based Video Information Processing System*. COMPSAC.

[14]

