



AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE
WYDZIAŁ INFORMATYKI, ELEKTRONIKI I TELEKOMUNIKACJI

KATEDRA TELEKOMUNIKACJI

Praca dyplomowa magisterska

Opracowanie, analiza i ocena algorytmu wyznaczania wartości MOS
Develope an Algorithm Predicting MOS as a Function of FR Metrics

Autor:

Kierunek studiów:

Opiekun pracy:

Pola Łącz

Teleinformatyka

dr hab. Lucjan Janowski

Kraków, 2019

Uprzedzony o odpowiedzialności karnej na podstawie art. 115 ust. 1 i 2 ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (t.j. Dz.U. z 2006 r. Nr 90, poz. 631 z późn. zm.): „Kto przywłaszcza sobie autorstwo albo wprowadza w błąd co do autorstwa całości lub części cudzego utworu albo artystycznego wykonania, podlega grzywnie, karze ograniczenia wolności albo pozbawienia wolności do lat 3. Tej samej karze podlega, kto rozpowszechnia bez podania nazwiska lub pseudonimu twórcy cudzy utwór w wersji oryginalnej albo w postaci opracowania, artystycznego wykonania albo publicznie zniekształca taki utwór, artystyczne wykonanie, fonogram, wideogram lub nadanie.”, a także uprzedzony o odpowiedzialności dyscyplinarnej na podstawie art. 211 ust. 1 ustawy z dnia 27 lipca 2005 r. Prawo o szkolnictwie wyższym (t.j. Dz. U. z 2012 r. poz. 572, z późn. zm.): „Za naruszenie przepisów obowiązujących w uczelni oraz za czyny uchybiające godności studenta student ponosi odpowiedzialność dyscyplinarną przed komisją dyscyplinarną albo przed sądem koleżeńskim samorządu studenckiego, zwanym dalej «sądem koleżeńskim».”, oświadczam, że niniejszą pracę dyplomową wykonałem(-am) osobiście i samodzielnie i że nie korzystałem(-am) ze źródeł innych niż wymienione w pracy.

Serdecznie dziękuję ... tu ciąg dalszych podziękowań np. dla promotora, żony, sąsiada itp.

Spis treści

1. Wstęp	7
2. Wprowadzenie teorytyczne	9
2.1. Cechy statystyczne wideo	9
2.2. Algorytmy uczenia maszynowego	10
3. Metodologia badań	11
3.1. Dane	11
3.2. Modele	11
4. Analiza i wnioski	13
5. Podsumowanie	15

1. Wstęp

Pierwsze obrazy wideo powstały już na początku XX wieku i opierały się na mechanicznie obracających się dyskach. Technologia ta istniała głównie w sferze badań akademickich i nie zdominowała rynku. Dopiero z wprowadzeniem cathode-ray tube (CRT), wraz z telewizją analogową, wideo zaczęło być wykorzystywane komercyjnie. Z czasem rozwój technologii pozwolił na wprowadzenie telewizji cyfrowej, która zapewniała wyższą jakość obrazu oraz lepsze wykorzystanie zasobów. Wideo razem z audio okazały się również znakomitym środkiem wymiany informacji. Coraz częściej wykorzystywane do komunikacji w czasie rzeczywistym zastępując tradycyjne połączenie telefoniczne w biznesie oraz dla zwykłych użytkowników. Również rozwój na rynku telefonów wspomógł powszechność wideo. W momencie kiedy praktycznie każdy aparat zaczął posiadać kamerę, wideo zaczęło konkurować ze zdjęciami jako metoda na utrwalenia danej chwili. Codziennie tak rejestrowane obrazy są przekazywane do rodziny, znajomych oddalonych o tysiące kilometrów. Kolejnym przykładem kiedy wideo zastępuje tradycyjne formy przekazu są blogi internetowe do tej pory prowadzone na zasadzie artykułów/postów, teraz zaczęły wykorzystywać wideo jako metodę przekazu informacji.

Dzięki coraz większym przepustowości i szerokiemu dostępowi do Internetu w najnowszych czasach wykreaował się jeszcze inny trend sprawiający że obrazy wideo są bardziej popularne. Mowa tu o platformach streamingowych takich jak - YouTube, Netflix czy HBOgo. Pozwalają one użytkownikom na oglądanie od krótkich filmików, przez serie, po pełnometrażowe filmy nawet w rozdzielczościach 4k.

Wszystkie wymienione wyżej aspekty sprawiły, że wideo stało się codziennością w życiu większości ludzi.

Na obecnym etapie rozwoju technologii, oczekiwania odbiorcy co do jakości otrzymywanego wideo znacznie wzrosły. Na drugiej szali pozostają ograniczenia dotyczące medium i optymalnego wykorzystania zasobów po stronie klienta i serwera. Odnosząc się do powyższego istotną kwestią staje się monitorowanie jakości transmitowanego wideo i dostosowywanie go do potrzeb użytkownika. Jednak problem w ocenie jakości wideo jest tu o tyle trudny, że dotychczas najbardziej wiarygodnym wskaźnikiem jest tu opinia ludzka, nie powiązana(?żadnym algorytmem?) z technicznymi aspektami ?obrazu?. W niniejszej Pracy zostanie przedstawiony algorytm pozwalający na bardziej zautomatyzowaną ocenę jakości wideo w oparciu o metryki full-reference (FR) i no-reference (NR) oraz zaprezentowana zostanie wykorzystana metodologia badań.

2. Wprowadzenie teoretyczne

W niniejszym rozdziale przedstawiono najważniejsze z zagadnień dotyczące przeprowadzonych badań. Opisane one zostały w sposób pozwalający czytelnikowi na odpowiednie zrozumienie dalszej części pracy, pomijając niezwiązane szczegóły. Pierwsza część rozdziału dotyczy tematu wideo. Przedstawiono jego definicję, oraz wybrane cechy statystyczne biorące udział w trakcie badań. W kolejnej części przedstawiono zagadnienia z obszaru uczenia maszynowego. Wyjaśniona została jej ogólna koncepcja, a następnie opisano użyte algorytmy.

2.1. Cechy statystyczne wideo

<definicja ciężko znaleźć :(> Wideo jest formą elektronicznego zapisu sygnału wizji (analogowego bądź cyfrowego). W swojej surowej postaci jest to sekwencja pojedynczych ramek. Takie pliki zajmują bardzo dużo przestrzeni dyskowej. Przykładowo obraz o rozdzielczości 1920x1080 z 24 ramkami na sekundę o długości 30 sekund zajmuje ponad 10 GB w pamięci komputera. Tak duże rozmiary znacząco ograniczają możliwości przechowywania i transmisji dla zwykłych użytkowników. Dlatego praktycznie każdy plik wideo wykorzystuje kodeki, czyli pewne ustandaryzowane zasady kompresji/dekompresji. Do najpopularniejszych należą H.265 i H.264 powszechnie używany w Internecie do transmitowania multimediów [1]. Dodatkowo wideo można charakteryzować poprzez wiele innych wskaźników. W poniżej zostały opisane te cechy statystyczne, które brały udział w badaniu.

- Rozdzielczość - miara określająca rozmiar ramki. J jednostką są pixele. Podawana jest zazwyczaj w następujący sposób: szerokość x wysokość. Do badań zostały użyte wideo po rozdzielczościach: 3840x2160, 1920x1080, 704x576, 640x480, 352x288.
- Klatki na sekundę(ang.frames per second, fps) - liczba ramek wyświetlonych w czasie sekundy. W telewizji jest to 25 ramek na sekundę. Do badań użyto: 60, 30, 25, 24 fps
- blockiness
- spatialact
- letterbox
- pillarbox

- blockloss
- blur
- temporalac
- blackout
- freezing
- exposure
- contrast
- brightness
- interlace
- noise
- slice
- flickering

opis metryk i ich wymienienie w oparciu o <https://ieeexplore.ieee.org/document/5506331>

- Ogólne informacje o wideo - czym jest, rodzaje.
- Przedstawienie wybranych cech statystycznych. (wszystkich?)

2.2. Algorytmy uczenia maszynowego

- Ogólne informacje uczeniu maszynowym/
- Przedstawienie wybranych algorytmów

3. Metodologia badań

3.1. Dane

- Wybrane narzędzia
- Opis zebranych danych
- Przedstawienie data flow(pobieranie-> czyszczenie->normalizacja->przygotowanie formatu dla modeli).
- Wizualizacja danych

3.2. Modele

- Opis zastosowanych parametrów/technik podczas trenowania.
- Przedstawienie wyników

4. Analiza i wnioski

- Interpretacja wyników
- Opis innych czynników mogących zaburzyć ich prawdziwość
- Co nie zostało uwzględnione

5. Podsumowanie

- Czy cel pracy został osiągnięty.
- Możliwości rozbudowy

Bibliografia

- [1] Anthony Romero. *On Line Alvis Manual*.

<https://video.ibm.com/blog/streaming-video-tips/what-is-video-encoding-code>
2018.