

**AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE**

**WYDZIAŁ INFORMATYKI, ELEKTRONIKI I TELEKOMUNIKACJIJ**

KATEDRA TELEKOMUNIKACJI

Praca dyplomowa magisterska

*Opracowanie, analiza i ocena algorytmu wyznaczania wartosci MOS´*

*Develope an Algorithm Predicting MOS as a Function of FR Metrics*

|  |  |
| --- | --- |
| Autor: | *Pola Ła˛cz* |
| Kierunek studiów: | *Teleinformatyka* |
| Opiekun pracy: | *dr hab. Lucjan Janowski* |

Kraków, 2019

*Uprzedzony o odpowiedzialnosci karnej na podstawie art. 115 ust. 1 i 2 ustawy z dnia 4 lu-´ tego 1994 r. o prawie autorskim i prawach pokrewnych (t.j. Dz.U. z 2006 r. Nr 90, poz. 631 z póz´n. zm.): „Kto przywłaszcza sobie autorstwo albo wprowadza w bła˛d co do autorstwa całosci lub cze˛´ sci cudzego utworu albo artystycznego wykonania, podlega grzywnie, karze ogra-´ niczenia wolnosci albo pozbawienia wolno´ sci do lat 3. Tej samej karze podlega, kto rozpo-´ wszechnia 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 takze uprzedzony o odpowie-˙ dzialnosci dyscyplinarnej na podstawie art. 211 ust. 1 ustawy z dnia 27 lipca 2005 r. Prawo´ o szkolnictwie wyzszym (t.j. Dz. U. z 2012 r. poz. 572, z póz´n. zm.): „Za naruszenie przepisów˙ obowia˛zuja˛cych w uczelni oraz za czyny uchybiaja˛ce godnosci studenta student ponosi odpo-´ wiedzialnosc´ dyscyplinarna˛przed komisja˛dyscyplinarna˛albo przed sa˛dem kole´ zen´skim samo-˙ rza˛du studenckiego, zwanym dalej «sa˛dem kolezen´skim».”, o˙ swiadczam,´ ze niniejsza˛prace˛ dy-˙ plomowa˛ wykonałem(-am) osobiscie i samodzielnie i´ ze nie korzystałem(-am) ze z´ródeł innych˙ niz wymienione w pracy.˙*

*Serdecznie dzie˛kuje˛ ...tu cia˛g dalszych podzie˛kowan´ np. dla promotora, zony, sa˛siada itp.˙*

Spis tresci´

[1. Wste˛p 7](#_Toc12727)

[2. Wprowadzenie teorytyczne 9](#_Toc12728)

[2.1. Wideo 9](#_Toc12729)

[2.2. Metryki 9](#_Toc12730)

[2.2.1. Metryki no-reference 10](#_Toc12731)

[2.2.2. Metryki full-reference 15](#_Toc12732)

[2.3. Algorytmy uczenia maszynowego 16](#_Toc12733)

[2.3.1. rrr 17](#_Toc12734)

[3. Metodologia badan´ 19](#_Toc12735)

[3.1. Dane 19](#_Toc12736)

[3.2. Modele 19](#_Toc12737)

[4. Analiza i wnioski 21](#_Toc12738)

[5. Podsumowanie 23](#_Toc12739)

*SPIS TRESCI´*

# Wste˛p

Pierwsze obrazy wideo powstały juz na pocza˛tku XX wieku i opierały sie˛ na mechanicznie obra-˙ caja˛cych sie˛ dyskach. Technologia ta istniała głownie w sferze badan akademickich i nie zdominowała´ rynku. Dopiero z wprowadzeniem cathode-ray tube (CRT), wraz z telewizja˛ analogowa˛, wideo zacze˛ło byc wykorzystywane komercyjnie. Z czasem rozwój technologii pozwolił na wprowadzenie telewizji´ cyfrowej, która zapewniała wyzsza˛ jako˙ s´c obrazu oraz lepsze wykorzystanie zasobów. Wideo razem z´ audio okazały sie˛ równiez znakomitym˙ srodkiem wymiany informacji. Coraz cze˛´ sciej wykorzystywane´ do komunikacji w czasie rzeczywistym zaste˛puja˛c tradycyjne poła˛czenie telefoniczne w biznesie oraz dla zwykłych uzytkowników. Równie˙ z rozwój na rynku telefonów wspomógł powszechno˙ s´c wideo. W´ momencie kiedy praktycznie kazdy aparat zacza˛ł posiada˙ c kamere˛, wideo zacze˛ło konkurowa´ c ze zdje˛-´ ciami jako metoda na utrwalenia danej chwili. Codziennie tak rejestrowane obrazy sa˛ przekazywane do rodziny, znajomych oddalonych o tysia˛ce kilometrów. Kolejnym przykładem kiedy wideo zaste˛puje tradycyjne formy przekazu sa˛ blogi internetowe do tej pory prowadzone na zasadzie artykułów/postów, teraz zacze˛ły wykorzystywac wideo jako metode˛ przekazu informacji.´

Dzie˛ki coraz wie˛kszym przepustowoscia˛i szerokiemu doste˛pu do Internetu w najnowszych czasach´ wykreował sie˛ jeszcze inny trend sprawiaja˛cy ze obrazy wideo sa˛ bardziej popularne. Mowa tu o plat-˙ formach streamingowych takich jak - YouTube, Netflix czy HBOgo. Pozwalaja˛ one uzytkownikom na˙ ogla˛danie od krótkich filmików, przez seriale, po pełnometrazowe filmy nawet w rozdzielczo˙ sciach 4k.´ Przewidywane jest, ze do 2022, a˙ z 82 procent całego ruchu IP, to be˛dzie wideo [1].˙

Wszystkie wymienione wyzej aspekty sprawiły,˙ ze wideo stało sie˛ codzienno˙ scia˛w´ zyciu wie˛kszo˙ sci´ ludzi.

Na obecnym etapie rozwoju technologii, oczekiwania odbiorcy co do jakosci otrzymywanego wideo´ znacznie wzrosły. Na drugiej szali pozostaja˛ograniczenia dotycza˛ce medium i optymalnego wykorzystania zasobów po stronie klienta i serwera(?dostawcy?). Odnosza˛c sie˛ do powyzszego istotna˛kwestia˛staje˙ sie˛ monitorowanie jakosci transmitowanego wideo i dostosowywanie go do potrzeb u´ zytkownika. Jednak˙ problem w ocenie jakosci wideo jest tu o tyle trudny,´ ze dotychczas najbardziej wiarygodnym wskaz´ni-˙ kiem jest tu opinia ludzka, nie powia˛zana(?zadnym agorytmem?) z technicznymi aspektami ?obrazu?˙ .

, cie˛zka do powia˛˙

W niniejszej Pracy zostanie przedstawiony algorytm pozwalaja˛cy na bardziej zautomatyzowana˛ ocene˛ jakosci wideo w oparciu o metryki full-reference (FR) i no-reference (NR) oraz zaprezentowana zostanie´ wykorzystana metodologia badan.´

cechami wideo

# Wprowadzenie teorytyczne

W niniejszym rozdziale przedstawiono najwazniejsze zagadnienia dla przeprowadzonych bada˙ n.´ Opisane one zostały w sposób pozwalaja˛cy czytelnikowi na odpowiednie zrozumienie dalszej cze˛ci pracy, pomijaja˛c niezwia˛zane szczegóły. Pierwsza cze˛s´c rozdziału dotyczy tematu wideo. Przedstawiono´ jego definicje˛, oraz wybrane cechy statystyczne biora˛ce udział w trakcie badan .W kolejnej cze˛´ sci przed-´ stawiono zagadnienia z obszaru uczenia maszynowego. Wyjasniona została jej ogólna koncepcja, a na-´ ste˛pnie opisano uzyte algorytmy.˙

## Wideo

<definicja cie˛zko znalez´˙ c :( > Wideo jest forma˛elektronicznego zapisu sygnału wizji (analogowego´ ba˛dz´ cyfrowego). W swojej surowej postaci jest to sekwencja pojedynczych ramek. Ramki natomiast składaja˛sie˛ z pixeli, a one sa˛definiowane poprez: luminance(*ang. luminace*) i chrominancje (*ang. chrominance*). Luminancja odpowiada za intensywnoc obrazu, a chrominacja za kolory.´

Surowe wideo zajmuje bardzo duzo przestrzeni dyskowej. Przykładowo obraz o rozdzielczo˙ sci´

1920x1080 z 24 ramkami na sekunde˛ o długosci 30 sekund zajmuje ponad´ 20 GB w pamie˛ci komputera. sprawdzic ile dok´ Tak duze rozmiary znacza˛co ograniczaja˛ mo˙ zliwo˙ sci przechowywania i transmisji dla zwykłych u´ zyt-˙ kowników. Dlatego praktycznie kazdy plik wideo wykorzystuje kodeki, czyli pewne ustandaryzowane˙ zasady kompresji/dekompresji. Do najpopularniejszych naleza˛ H.265 i H.264 powszechnie u˙ zywany w˙ Internecie do transmitowania multimediów [2]. Błe˛dy podczas tego procesu, transmisji czy nagrywania moga˛ powodowac powstanie zaburze´ n w odtwarzanym obrazie. Aby mierzy´ c poziom uszkodze´ n oraz´ monitorowac stan wideo badacze zdefiniowali wiele metryk.´

.

## Metryki

Jakos´c danego wideo jest najtrafniej oceniona dzie˛ki tak zwanej ocenie subiektywnej. Polega ona na´ okresleniu jako´ sci odbioru przez człowieka na podstawie jego odczu´ c. Niestety aby takie badanie było´ miarodajne nalezy je przeprowadzi˙ c w odpowiednich warunkach, przygotowa´ c zestaw testów, zrekruto-´ wac uczestników oraz cały czas nadzorowa´ c jego przebieg. Wszystko to generuje du´ ze koszty i zabiera˙

cenny czas, ale przede wszystkim nie pozwala na ocene˛ jakosci w czasie rzeczywistym na przykład pod-´ czas streamowania video. Problem ten rozwia˛zuja˛ metryki jakosci wideo. które w przeciwie´ nstwie do´ oceny subiektywnej, bazuja˛na obiektywnych pomiarach własciwo´ sci[3].´

Metryki video moga˛ zostac sklasyfikowane na podstawie tego czy do ich policzenia wymagana jest´ obecnos´c niezakłóconego pliku video. Metryki, które wymagaja˛takiego pliku okre´ slane sa˛mianem´ *ang. full-reference* (FR). Porównuja˛one własciwo´ sci zakłóconego i nie zakłóconego video, aby na podstawie´ tak otrzymanych informacji dokonac estymaty jako´ sci. Innym rodzajem metryk sa˛ metryki´ *ang. noreference* (NR). Dokonuja˛ one oceny jakosci na podstawie˛ zniekształconego video i nie wymagaja˛ do´ tego zadnych dodatkowych informacji[3].˙

W ponizszych podrozdziałach w po krótce zostały opisane metryki typu NR i FR.˙

### Metryki no-reference

* Blokowos´c (´ *ang. Blockiness*) – powstaje podczas procesu kwantyzacji bloków pikseli i objawia sie˛ poprzez zauwazalna˛granice mie˛dzy tymi blokami[4]. Skala przyje˛ta w pracy: 0-3570. Im wie˛ksza˙ wartos´c, tym mniej widoczne zakłócenie. Dla obrazu bez zakłóce´ n około 0.9-1.01[5]´



Rys. 2.1. Przykładowy obraz z zakłóceniami spowodowanymi blokowoscia˛[5]´

* Aktywnos´c Przestrzenna (´ *ang. Spatial Activity*) – opisuje połozenie obiektu na ramce oraz jego˙ relacje z innymi obiektami. Pozwala rozrózni˙ c czy aktywno´ s´c na danym wideo jest czynno´ scia˛´ statyczna˛(osoba wykonuja˛ca czynnos´c pozostaje w jednym miejscu) czy mobilna˛(osoba wykonu-´ ja˛ca czynnos´c porusza sie˛ wzdłu´ z pola widzenia)[6]. Skala przyje˛ta w pracy: 0-270. Im wie˛ksza˙ wartos´c tym wie˛ksza aktywno´ s´c przestrzenna[5]´



Rys. 2.2. Przykład detekcji obiektów w ramce, niezbe˛dne do otrzymania informacji o aktywnosci przestrzennej[5]

* Letterbox i Pillarbox - letterbox jest technika˛ pozwalaja˛ca˛ na wyswietlanie obrazów o wy´ zszym˙ współczynniku proporcji na odbiornikach o mniejszym współczynniku. Polega ona na dołozeniu˙ dwóch czarnych pasów na górze i dole ramki.[7]. Natomiast pillarbox uzywane jest przy wy˙ swie-´ tlaniu obrazu w mniejszym współczynniku proporcji na ekranie o wie˛kszym . Do obrazu dodawane sa˛wtedy pasy po obu bokach ekranu. Skala przyje˛ta w pracy 0-1.



(a) opis[5] (b) opis[5]

* Straty bloków (*ang. Blockloss*) – wskaz´nik informuja˛cy o ilosci brakuja˛cych bloków(u´ zywanych˙ podczas DCT(Dyskretna transformacja kosinusowa)). Skala przyje˛ta w pracy: 0-200.Im wie˛ksza wartos´c tym wie˛ksze zniekształcenie obrazu.´



Rys. 2.4. opis[5]

* Rozmycie (*ang. Blur*) – Istnieje wiele rodzajów rozmycia, miedzy innymi te niepoza˛dane, wy-˙ ste˛puja˛ce po przeprowadzeniu kompresji ba˛dz´ przez poruszenie kamerypodczas nagrywania. Gdy wyste˛puje ten efekt obiekty na obrazie traca˛ostros´c krawe˛dzi.[8]. Skala przyjeta w pracy: 0-70.´



Rys. 2.5. A non-floating figure *with* a caption!

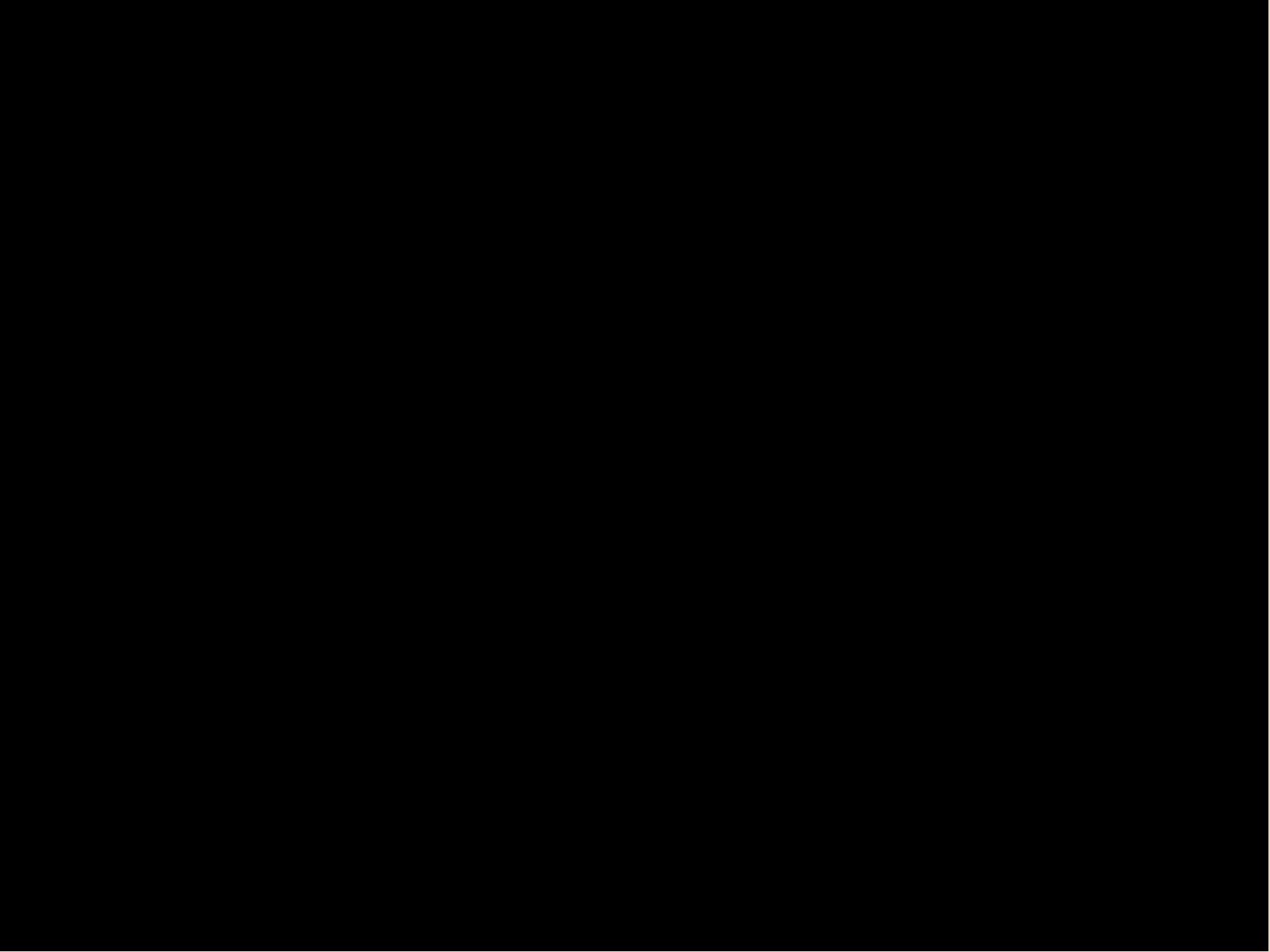
* Aktywnsc czasowa (´ *ang. Temporal Activity*) – opisuje intensuwnosc ruchu obiektów w czasie.´

Cze˛sto ła˛czone z wskaz´nikiem Aktywnosci przestrzenej. Skala przyje˛ta w pracy: 0-255.´



Rys. 2.6. A non-floating figure *with* a caption!

* Wyciemnienia (*ang. Blackout*) – opisuje zjawisko kiedy obraz całkowicie zanika[9]. Skala przyje˛ta w pracy : 0-1



Rys. 2.7. A non-floating figure *with* a caption!

elacpo

* Zamrozenie (˙ *ang. Freezing*) – Informuje o efekcie czasowego zatrzymania zatrzymania obrazu, które powoduje odczucie "przycie˛cia"wideo[10]. Skala przyje˛ta w pracy:0-1.
* Ekspozycja (*ang. Exposure*) – ten typ zakłócen jest spowodowany niezbilansowana˛ jasno´ scia˛ ra-´ mek. Widz ma odczucie ze wideo jest zbyt ciemna lub zbyt jasne[10]. Skala przyje˛ta w badaniach 0-255.



Rys. 2.8. A non-floating figure *with* a caption!

* Kontrast (*ang. Contrast*) – Opisuje róznice pomie˛dzy jasnymi, a ciemnymi obszarami obrazu.˙ Skala przyje˛ta w pracy: 0-120.



Rys. 2.9. A non-floating figure *with* a caption!

* Jasnosc (´ *ang. Brightness*) – jest powia˛zana z problemem Ekspozycji. Jej zbyt duza warto˙ s´c mo´ ze˙ skutkowac odczuciem przeswietlenia ramki. Skala przyje˛ta w pracy: 0-1.´



Rys. 2.10. A non-floating figure *with* a caption!

* Szum (*ang. noise*) – jest to rodzaj zaburzenia obrazu spowodowany wyste˛powaniem niekontolowanych wzorców dla intensywnosci wy´ swietlania pixeli[10]. Skala przyje˛ta w pracy: 0-120.´



Rys. 2.11. A non-floating figure *with* a caption!

* Pocie˛cie (*ang. Slice*) – objawia sie˛ efektem niepasuja˛cych do całosci poziomych pasów. Zwia˛zane´ jest to utrata˛pakietów danych podczas transmisji wideo. NIE DZIALA POPRAWNIE !



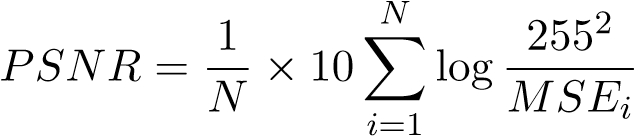
Rys. 2.12. A non-floating figure *with* a caption!

Oprócz wyzej wymienionych jako dane treningowe zostały równie˙ z u˙ zyte cechy wideo takie jak:˙

* Rozdzielczos´c (´ *Resolution*) – miara okreslaja˛ca rozmiar ramki. Jednostka˛ sa˛ pixele. Poda-´ wana jest zazwyczaj w naste˛puja˛cy sposób: szerokos´c x wysoko´ s´c. Do bada´ n zostały u´ zyto˙ wideo o rozdzielczosciach: 3840x2160, 1920x1080, 704x576, 640x480, 352x288.´
* Klatki na sekunde˛ (*ang. fps, frames per second*) – liczba ramek wyswietlonych w czasie´ sekundy. W telewizji jest to 25 ramek na sekunde˛. Do badan u´ zyto: 60, 30, 25, 24 fps˙
* Kodowanie kolorów (*ang. Color encoding*) – todo

### Metryki full-reference

* PSNR (*ang. Peak Signal-to-Noise Ratio*) – prosta do policzenia i równiez cze˛sto stosowane z˙ metryk typu FR. Przykładowa implementacja dla PSNR 2.1 (gdzie MSE jest błe˛dem srednio´ kwadratowym dla ramki). W pierwszej kolejnosci liczone jest zniekształcenie dla wszystkich´ ramek w stosunku do ich oryginału. Zazwyczaj do obliczan wykorzystuje sie˛ informacje˛ o´ luminacji pixela. Póz´niej naste˛puje usrednienie wyniku[11].´

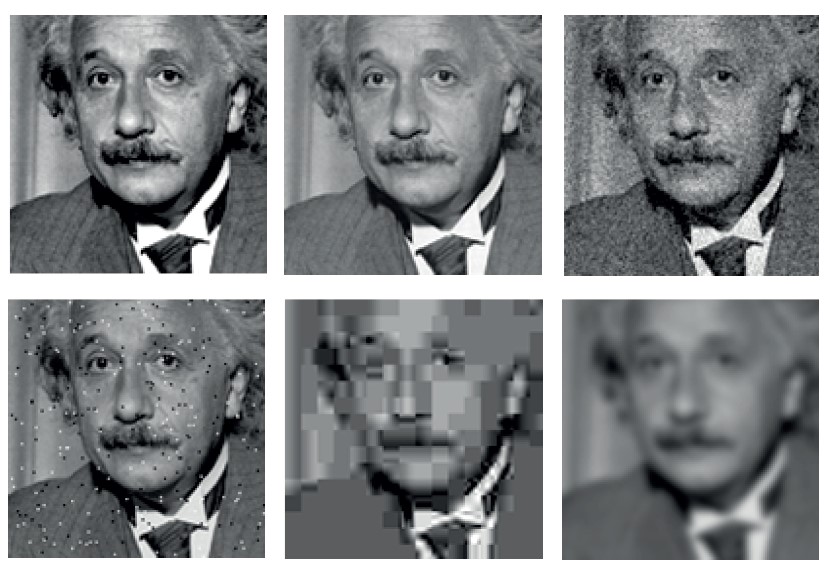
 (2.1)

Mimo zalet PSNR nie zawsze spełnia swoje zadanie. Ilustracja 2.13 przedstawia w rózny˙ sposób zniekształcone obrazy. Kazdy z nich otrzymał ta˛ sama˛ ocene˛ PSNR, a jednak ocena˙ subiektywna wskazuje na rózny poziom jako˙ sci.´

* SSIM (*ang. Structural Similarity Index*) – jest jedna z najcze˛sciej stosowanych metryk oceny´ jakosci wideo. Równie´ z, jak poprzednia metryka, jest ona prosta w obliczeniach. Uwzgle˛dnia˙ w swojej funkcji luminacje, kontrast i strukture˛ obu 2.2. Gdzie S jest blokiem pixeli z ramki z oryginalnej sekwencji wideo, a S’ jej zniekształcona˛ wersja˛. *SSIM* ∈*<* 0*,*1 *>*, gdzie 1 oznacza brak zniekształcen, a 0 idealne niedopasowanie [12][13].´

#### SSIM = I(S,S0)c(S,S0)s(S,S0) (2.2)

*2.3. Algorytmy uczenia maszynowego*



Rys. 2.13. opis

* MS-SSIM (*ang. multi-scale SSIM*) – jest modyfikacja˛ SSIM’a. Polega na próbkowaniu obrazu, w kazdej skali, przed zastosowaniem SSIM’a, a po przeprowadzeniu oblicze˙ n, ka´ zda˙ z skali jest ła˛czona, z odpowiednimi wagami, w jedna˛ całosc. Zabieg ma na celu poprawe´ trafnosci oceny jakosci[13].
* VMAF (*ang. Video Multimethod Assessment Fusion*) – Metryka rozwijana przez Netflix we współpracy z University of Southern California. Jej działanie opiera sie˛ podła˛czaniu wyników kilku metryk podstawowych. VMAF nadaje kazdej z metryk odpowiednie wagi ,a dzieje˙ sie˛ to dzie˛ki zastosowaniu algorytmu uczenia maszynowego, w tym przypadku SVM regressor (*ang. Support Vector Machine regressor*). Takie rozwia˛zanie ma zapewnic wykorzystanie´ mocnych stron poszczególnych metryk z pominie˛ciem ich słabosci [14].´

## Algorytmy uczenia maszynowego

Obecne czasy charakteryzuja˛sie˛ ogromna˛iloscia˛danych, które zewsza˛d otaczaja˛człowieka i jego otoczenie. W kazdej sekundzie ró˙ znego rodzaju dane sa˛zbierane, przesyłane, prezentowane i ana-˙ lizowane. Aby móc w optymalny sposób wykorzystac drzemia˛cy w nich potencjał zostały opraco-´ wane pewen algorytmy, bazuja˛ce na statystyce. Potrafia˛one stworzyc modele matematyczne zada-´ nego zdarzenia i na ich podstawie dokonac predykcji, klasyfikacji czy zinterpretowa´ c i zareagowa´ c´ na zadane sytuacje(?uczyc sie na swich wlasnych obliczeniach?). Dziedzina nauki zajmuja˛ca sie˛ takimi algorytmami nazywana jest Uczeniem Maszynowym (*ang. Machine Learning*).

Wyrózniane sa˛dwa typy algorytmów uczenia maszynowego z uczeniem nadzorowanym (˙ *ang. supervised learning*) i bez nadzorowania(*ang. unsupervised learning*). Róznica wyste˛puje podczas˙ procesu tworzenia modelu – trenowania. Pierwsza grupa, aby poprawnie zbudeowac model wy-´ maga podania danych

*2.3. Algorytmy uczenia maszynowego*

### rrr

* Ogólne informacje uczeniu maszynowym/
* Przedstawienie wybranych algorytmów

*2.3. Algorytmy uczenia maszynowego*

# Metodologia badan´

## Dane

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

## Modele

* Opis zastosowanych paramtrów/technik podczas trenowania.
* Przedstawienie wyników

*3.2. Modele*

# Analiza i wnioski

* Interpretacja wyników
* Opis innych czynników moga˛cych zaburzyc ich prawdziwo´ c´
* Co nie zostało uwzgle˛dnione

# Podsumowanie

Zle nagranego wideo nie uratuje zadna metryka Metryki powinny dostosowywac sie do (rodzaju zakłócen)zaklocen´

* Czy cel pracy został osia˛gniety.
* Mozliwoci rozbudowy˙

Bibliografia

1. Cisco. *Forecast and Trends 2017-2022*.

https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual

2019.

1. Anthony Romero. *On Line Alvis Manual*.

https://video.ibm.com/blog/streaming-video-tips/what-is-video-encoding-cod

2018.

1. Christian Keimel. *Design of Video Quality Metrics with Multi-Way Data Analysis: A data driven approach*. Spraw. tech. 2016, s. 31–32.
2. M. Ghanbari M.T. Qadri K.T. Tan. *Frequency domain blockiness measurement for image quality assessment, Introduction*.

https://ieeexplore.ieee.org/document/5645871. 2010.

1. Katedra Telekomunikacjii AGH IEiT. *Video Quality Indicators*.

http://vq.kt.agh.edu.pl/index.html. 9999.

1. William Blackburn Shreeya Sengupta Hui Wang. *Spatial information in classification of activity videos*. Spraw. tech. School of Computing i Mathematics, University Of Ulster, 2015.
2. *16:9 aspect ratio conversion by letterbox method for an MPEG image*. 2000.
3. S. Winkler P. Marziliano F. Dufaux. *A no-reference perceptual blur metric*.

https://ieeexplore.ieee.org/document/5645871. 2002.

1. Abdelhamid Mellouki Hala Rifaï Samer Mohammed. *A brief synthesis of QoS-QoE methodologies*.

https://ieeexplore.ieee.org/document/5898880. 2011.

1. Mylène C. Q. Farias Mikołaj Leszczuk Mateusz Hanusiak. *Recent developments in visual quality monitoring by key performance indicators*. Video artefacts and related key performance indicators for automated quality checking. 2014.
2. A. Ebrahimi A. T. Nasrabadi M. A. Shirsavar. *Investigating the PSNR calculation methods for video sequences with source and channel distortions*. https://ieeexplore.ieee.org/document/6873482. 2014.

*BIBLIOGRAFIA*

1. xxx. *Design of Video Quality Metrics with Multi-way Data Analysis*. sciagniete na dysk. 2010.
2. xxx. *Image quality metrics: PSNR vs. SSIM*. sciagniete na dysk. 2010.
3. xxx. *Toward A Practical Perceptual Video Quality Metric*.

https://medium.com/netflix-techblog/toward-a-practical-perceptual-video-qualit

9999.