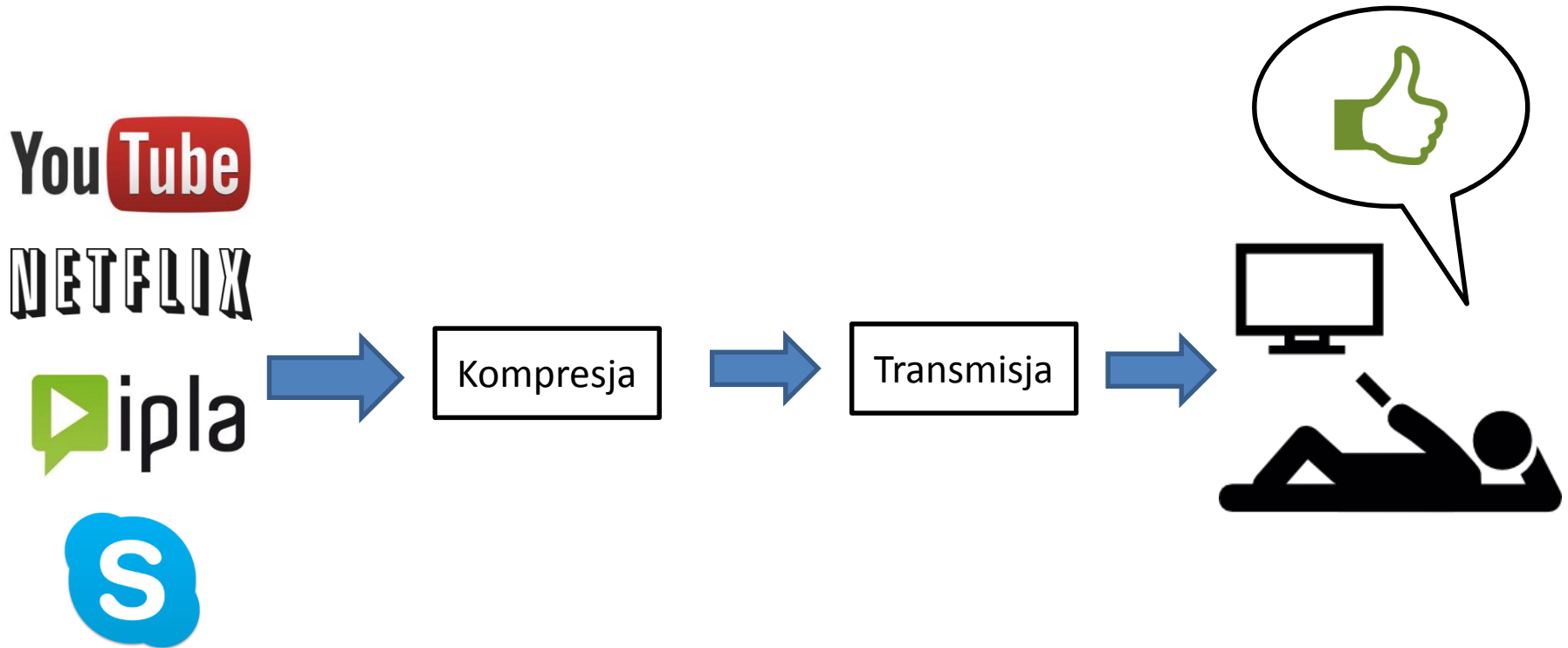


Systemy i Terminale Multimedialne

Obiektywne metody pomiaru jakości obrazu

Marcin Szykulski

Wprowadzenie



Jak osiągnąć kompromis przepływność/jakość?

Podstawowe informacje

- Jakość obrazu wideo można oceniać w sposób:
 - **Subiektywny** (metryki obserwacyjne) – psychowizualne testy oceny jakości przeprowadzane przy pomocy grona specjalistów według ustalonych reguł
 - **Obiektywny** (metryki obliczeniowe) – wielkości wyznaczane zgodnie z ustaloną zależnością, często optymalizowane z wykorzystaniem subiektywnych ocen specjalistów (podejście subiektywno-obiektywne)

Podstawowe informacje

- Video Quality Expert Group (VQEG)
 - Powołana w 1997 r.
 - Skupia ekspertów, zajmujących się zarówno subiektywną jak i obiektywną oceną jakości obrazu wideo
 - Głównym zadaniem ekspertów z VQEG jest planowanie i przeprowadzanie testów walidacji obiektywnych metod pomiaru jakości obrazu wideo – tworzenie norm w ramach sektorów ITU-T i ITU-R
 - Większość ekspertów jest zaangażowana w prace organizacji standaryzującej ITU
- Laboratory for Image and Video Engineering (LIVE)
 - Zespół University of Texas zajmujący się perceptualną analizą obrazu
 - Autorzy m.in. obiektywnej metryki SSIM, szeregu publicznie dostępnych korpusów zawierających wideo i obrazy ocenione subiektywnie przez grono specjalistów

Subiektywne badanie jakości obrazu

- Subiektywna ocena jakości obrazu wizyjnego:
 - Metodykę prowadzenia badań w celu dokonania subiektywnej oceny jakości opisano w normach ITU-R BT.500 (TV) oraz ITU-T P.910 (multimedia)
 - Przygotowanie pomieszczenia i sprzętu testowego zgodnie z normami (oświetlenie pomieszczenia, kalibracja monitora, maksymalny kąt obserwacji)
 - Wybór próbek testowych – powinny zawierać zniekształcenia o różnym nasileniu
 - Przeprowadzenie sformalizowanej projekcji badanych próbek wideo z udziałem odpowiedniej (w sensie statystycznym) liczby osób
 - Jednorazowe sesje testowe powinny trwać nie dłużej niż 30 min.
 - Miernikiem jakości obrazu jest zazwyczaj parametr MOS (ang. Mean Opinion Score)

Ocena	Jakość	Zakłócenia
5	doskonała	nieobecne
4	dobra	widoczne ale nieuciążliwe
3	przeciętna	lekko uciążliwe
2	niska	uciążliwe
1	zła	bardzo uciążliwe

Subiektywne badanie jakości obrazu

Metody subiektywne są długotrwałe i kosztowne, stąd potrzeba opracowania metod dokonujących oceny jakości obrazu w sposób automatyczny.

Przykładowa konfiguracja pomieszczenia służącego do oglądania obrazu wideo w celu dokonania subiektywnej oceny jakości obrazu:



Sprzęt do odtwarzania/nadawania nieskompresowanego materiału w formacie HDTV

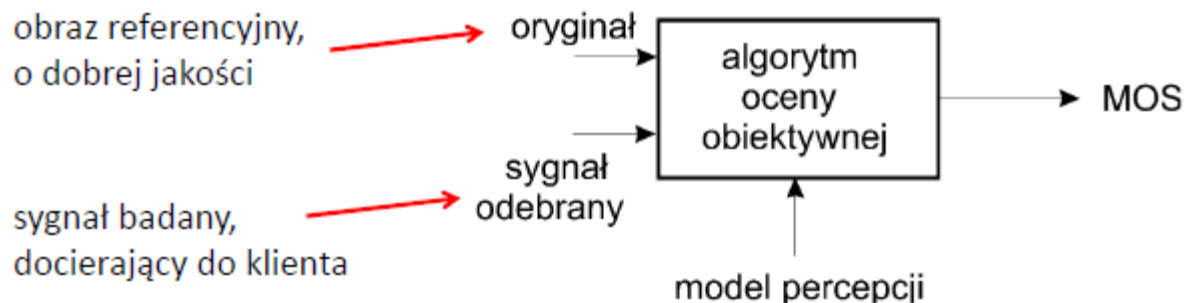


Pomieszczenie projekcyjne
(ang. viewing room)

Obiektywne badanie jakości obrazu

- Metody obiektywne opierają się na jednoznacznie mierzalnych wartościach – nie wymagają obecności ludzi, dla danego zestawu parametrów zawsze otrzyma się ten sam wynik
- Obiektywne metody oceny jakości zawierają model ludzkiej percepcji (HVS – ang. *human visual system*), w celu zbliżenia się do ocen wystawianych subiektywnie przez specjalistów. (Wyjątkiem – proste miary, takie jak np. PSNR)
- W HVS modelowaniu podlega m.in. bezwładność czasową oka, zmniejszoną czułość na chrominancję, zjawisko maskowania

Schemat ogólny obiektywnej metody badania jakości:



Rodzaje zniekształceń i zakłóceń obrazu wizyjnego

- Wynikające z kompresji:

- Efekty blokowe – zauważalne stają się granice bloków (np. 8x8 pikseli – wynik DCT)
- Zniekształcenia konturów (nieostrość, załamania, poruszanie się)
- Przekłamanie w odtworzeniu kolorów (zmiana odcieni, nasycenia)
- Szum kwantyzacji (np. „śnieżenie”)



- Wynikające z transmisji danych w pakietach:

- Zatrzymanie obrazu – gdy szereg kolejnych ramek zostaje niepoprawnie zakodowanych, zostaje podstawiona ostatnia poprawnie zakodowana ramka – wynik dużych zakłóceń w kanale transmisyjnym
- Chwilowy zanik obrazu lub jego części, często połączony z pojawieniem się na ekranie kolorowej tekstury – przy bardzo dużych zakłóceniach.

Wymienione czynniki muszą być uwzględnione przy opracowywaniu obiektywnego algorytmu wyliczania jakości sygnału wideo.

Rodzaje zniekształceń i zakłóceń obrazu wizyjnego



(a)



(b)



(c)



(d)

- (a) Kompresja MPEG-2
- (b) Kompresja h.264
- (c) Straty pakietów w sieci IP
- (d) Straty pakietów w sieci komórkowej

Rodzaje obiektywnych metryk jakości obrazu

- **Metryki względne (ang. *Full Reference Methods – FR*)**
 - Oparte na porównywaniu obrazów względem siebie – jeden z nich jest obrazem referencyjnym – powinien to być obraz oryginalny (nie poddany kompresji)
 - Do określenia obiektywnej jakości obrazu wykorzystuje się wynik porównania punkt po punkcie (każdy piksel oryginalnego obrazu z odpowiadającym pikselem obrazu zdegradowanego)
 - Najdokładniejsze, ale najbardziej złożone obliczeniowo
 - Ograniczona możliwość stosowania – wymagany jest dostęp do materiału źródłowego

Rodzaje obiektywnych metryk jakości obrazu

- **Metryki bezwzględne (ang. *No-Reference Methods* – NR)**
 - Oparte na analizie parametrów obrazu zdegradowanego
 - Wykorzystywane gdy oryginalny obraz nie jest dostępny
 - Wykorzystują przekształcenia matematyczne do obliczenia charakterystycznych cech obrazu (parametry przestrzenne) albo ich sekwencji (parametry czasowe)
 - Mniej dokładne niż metryki względne, ale łatwiejsze do obliczenia
- **Rodzaje metryk bezwzględnych:**
 - Oparte na pikselach – analiza jakości na podstawie informacji zawartej w pikselach obrazu
 - Oparte na parametrach strumienia – wykorzystują informacje z nagłówek pakietów (np. MPEG-TS), nie dokonują dekodowania strumienia wideo. Mało skuteczne, ale najłatwiejsze do obliczenia
 - Hybrydowe

Metryki względne

- **PAE (Peak Absolute Error)**

- Szczytowy błąd absolutny

$$PAE = \max_{i=1\dots N} |x_i - y_i|$$

- N – liczba pikseli w analizowanej klatce strumienia wideo
- x_i – i -ty piksel obrazu oryginalnego
- y_i – i -ty piksel obrazu zakłóconego (odebranego)

Metryki względne

- **MSE (Mean Square Error)**
 - Błąd średniokwadratowy wielkości pikseli obrazu

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2$$

- N – liczba pikseli w analizowanej klatce strumienia wideo
- x_i – i -ty piksel obrazu oryginalnego
- y_i – i -ty piksel obrazu zakłóconego (odebranego)

Metryki względne

- **MSE (Mean Square Error) – cd.**
 - Pozwala ocenić stopień rekonstrukcji obrazu przez dekodery
 - Wartości: od 0 (brak różnicy) do $255^2 = 65025$ (maksymalna różnica przy 8 bitowej głębi kolorów) – powinna być jak najmniejsza



Original

Processed

MSE

Metryki względne

- **PSNR (Peak Signal-to-Noise Ratio)**
 - Jedna z najczęściej używanych miar, opisuje stosunek sygnału szczytowego do szumu [dB]
 - Jest to błąd średniokwadratowy odniesiony do maksymalnej możliwej różnicy, wyrażony w skali logarytmicznej:

$$\text{PSNR} = 10 \log_{10} \frac{L^2}{\text{MSE}}$$

L – zakres wartości piksela

(np. gdy każdy piksel jest zapisany na 8 bitach, wartość parametru L będzie równa 255)

Metryki względne

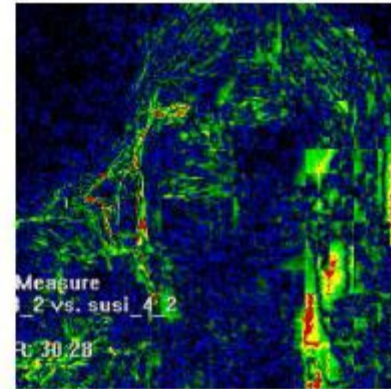
- **PSNR** – ciąg dalszy
 - Przyjmowane wartości dla 8 bitowego koloru: od 0 (maksymalna różnica – czerwony) do 50 dB - (brak różnic – czarny) – im większe PSNR, tym lepiej



Original



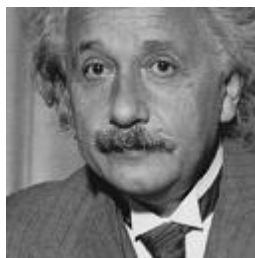
Processed



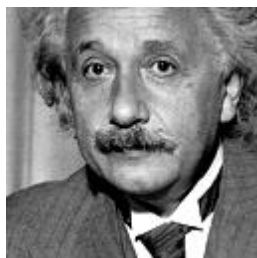
PSNR

Metryki względne

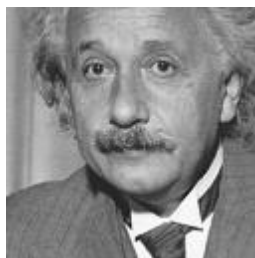
Parametry MSE i PSNR są często stosowane ze względu na łatwość wyznaczenia, jednak słabo odzwierciedlają wrażenia percepcyjne odbiorcy!



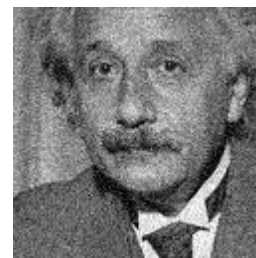
(a) Obraz oryginalny MSE=0



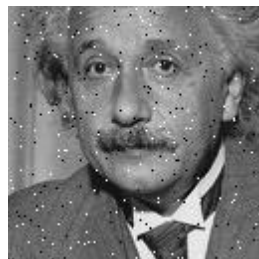
(b) MSE=306



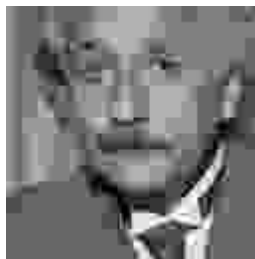
(c) MSE=309



(d) MSE=309



(e) MSE=313



(f) MSE=309



(g) MSE=308

Metryki względne

Parametry MSE i PSNR są często stosowane ze względu na łatwość wyznaczenia, jednak słabo odzwierciedlają wrażenia percepcyjne odbiorcy!



MSE = 27.10

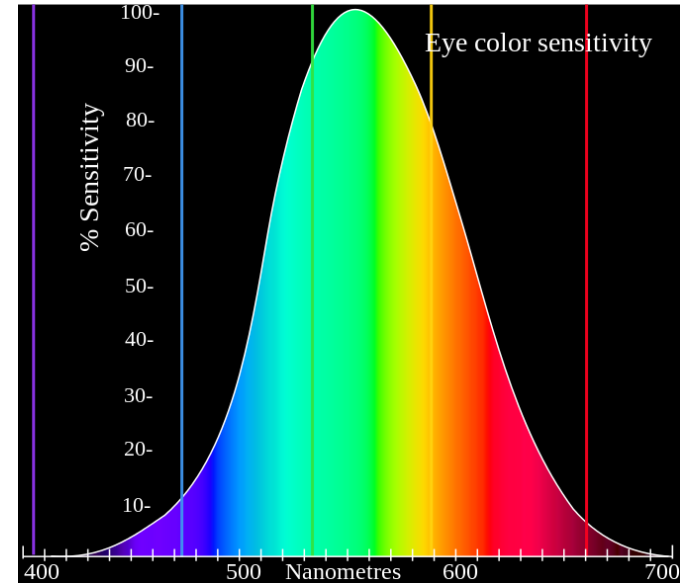


MSE = 21.26

Metryki względne

Zmysł wzroku jest szczególnie uwrażliwiony na kontrast:

- Uwypukła granice pomiędzy polami o niewiele różniące się luminancji
- Z drugiej strony – niejednorodne tło jest w stanie całkowicie zamaskować obiekt (staje się on niezauważalny)



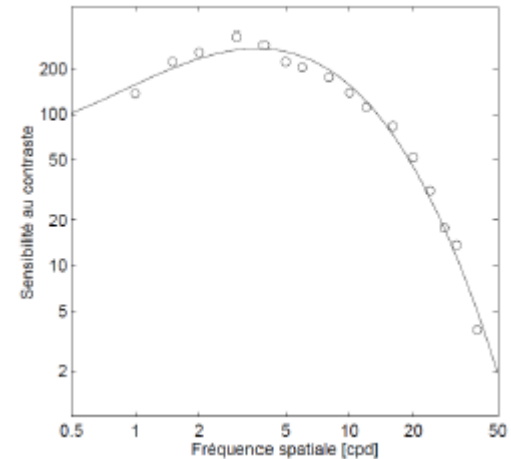
Ponadto, duże znaczenie w procesie oceny jakości obrazu mają inne cechy zmysłu wzroku:

- Wrażliwość na barwy (oko jest bardziej czułe na barwę zieloną, niż na czerwoną czy niebieską)
- Bezwładność (w dziedzinie czasu)
- Rozdzielczość kątowna

Metryki względne

Najważniejszy problem w modelowaniu ludzkiej percepcji (HVS - Human Visual System) – uwzględnienie wrażliwości na kontrast, oraz spadku tej wrażliwości w obecności niejednorodnego tła (maskowanie)

- Maksymalna wrażliwość na kontrast (minimalny próg detekcji) występuje w procesie obserwowania obiektów o rozmiarach ok. $\frac{1}{4}$ stopnia („częstotliwość przestrzenna” 4 cykle zmian luminancji / stopień)
- Dla mniejszych „częstotliwości przestrzennych” wrażliwość na kontrast (CSF – *contrast sensitivity function*, zdefiniowana jako odwrotność progu detekcji), powoli spada, a dla większych częstotliwości – spada bardzo szybko

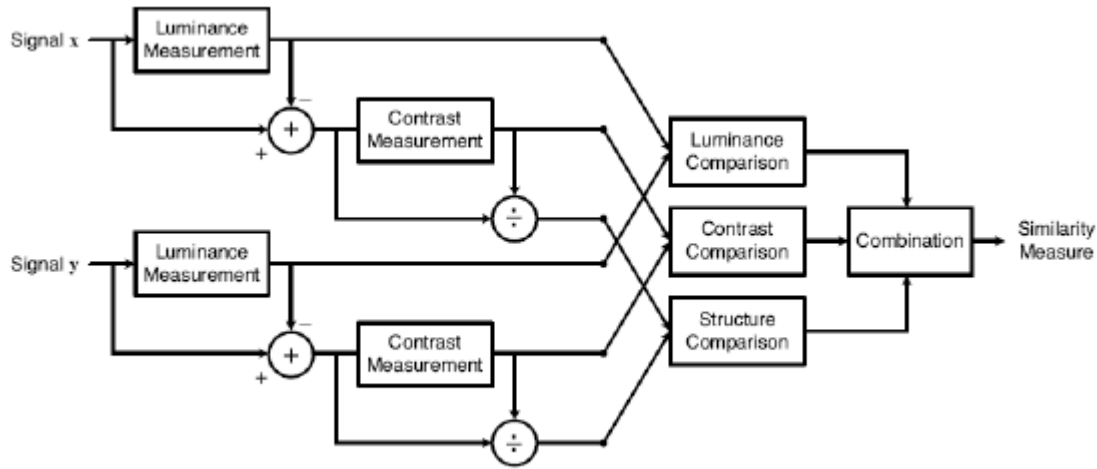


Rys. Wrażliwość na kontrast w funkcji częstotliwości przestrzennej

Metryki względne

- **SSIM (Structure Similarity Index Measure)**

- Pierwsza metryka, która porównuje nie wartości pikseli ale elementy obrazu spostrzegane przez człowieka
- Lepiej opisuje różnice jakości obrazu niż np. PSNR,
- Uwzględnia trzy typy zniekształceń: luminancji, kontrastu i struktury
- Końcowy indeks SSIM uwzględnia łącznie wszystkie te zniekształcenia,
- Przyjmuje wartości: od -1 (maksymalna różnica) do 1 (brak różnic), wyższa wartość to lepsza jakość.



Metryki względne

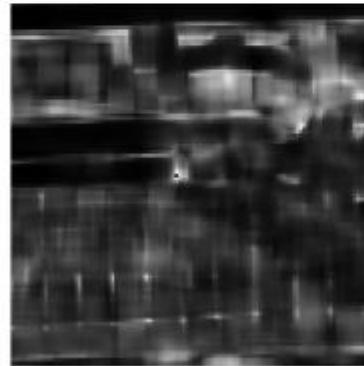
- **SSIM** - ciąg dalszy



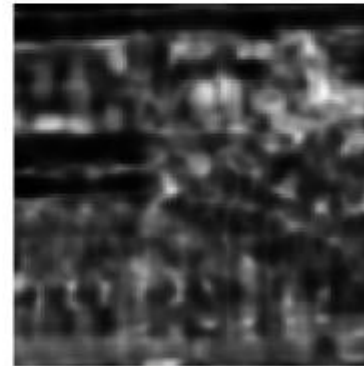
Original



Compressed



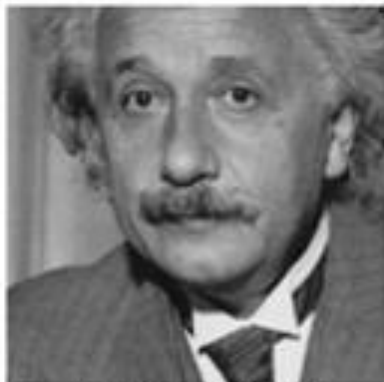
SSIM (fast) visualization



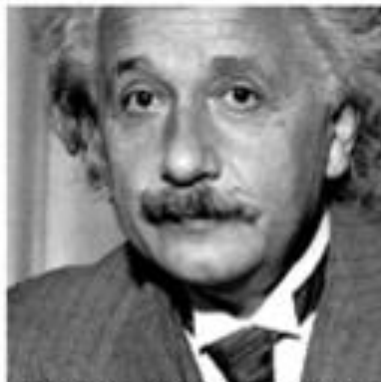
SSIM (precise) visualization

Metryki względne

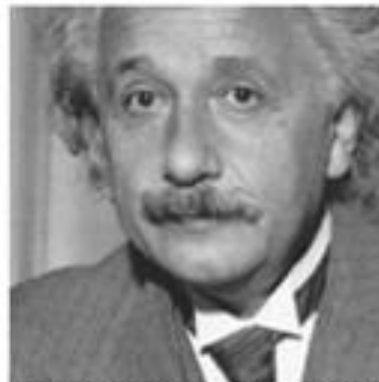
- **SSIM** - ciąg dalszy



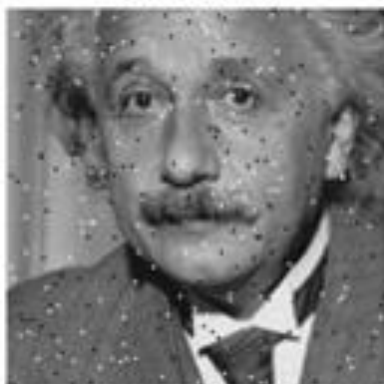
(a) MSE=0, SSIM=1
CW-SSIM=1



(b) MSE=306, SSIM=0.928
CW-SSIM=0.938



(c) MSE=309, SSIM=0.987
CW-SSIM=1.000



(e) MSE=313, SSIM=0.730
CW-SSIM=0.811



(f) MSE=309, SSIM=0.580
CW-SSIM=0.633



(g) MSE=308, SSIM=0.641
CW-SSIM=0.603

Metryki względne

- **VQM (Video Quality Measure)**

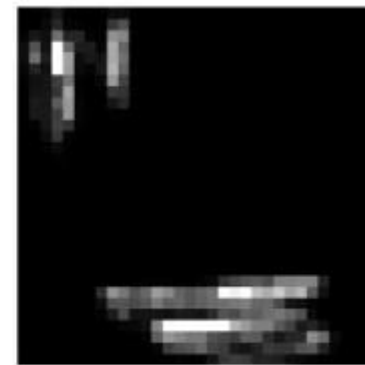
- Ma za zadanie ocenić jakość obrazu, czyli stopień zniekształceń widzianych przez człowieka
- Wartości VQM są w przybliżeniu skorelowane z subiektywnymi ocenami widzów
- Algorytm dokonuje operacji na współczynnikach transformaty kosinusowej DCT (obliczenie lokalnego kontrastu i porównanie z funkcją postrzegalności kontrastu)
- Wartości: 0 oznacza brak różnic, im większa wartość tym większe różnice (gorsza jakość).



Original



Processed



VQM

Metryki bezwzględne

Człowiek z reguły bardzo łatwo ocenia wielkość zniekształcenia obrazu nie oglądając niezniekształconego oryginału.

Trudno jest natomiast skonstruować program komputerowy, który bez dostępu do obrazu odniesienia poprawnie oszacowywałby jakość obrazu poddanego ocenie.

Wszystkie dotychczas opracowane metryki bezwzględne służą wyłącznie do pomiaru wielkości zakłóceń konkretnego typu.

Metryki bezwzględne

- **Blurring Metric**

- Miara ta opisuje zniekształcenia objawiające się rozmyciem (blur) szczegółów obrazu, zwłaszcza krawędzi obiektów (charakteryzujących się gwałtowną zmianą barwy lub jasności) – wykorzystywane są algorytmy detekcji krawędzi np. Canny’ego, Sobela
- Rozmycie szczegółów może być artefaktem kompresji (wynikającym z „uśredniania” bloków pikseli), może też wynikać z zastosowania układu rozmywającego zniekształcenia blokowe (deblocking), występującego w większości dekoderek

Metryki bezwzględne

- **Blurring Metric**

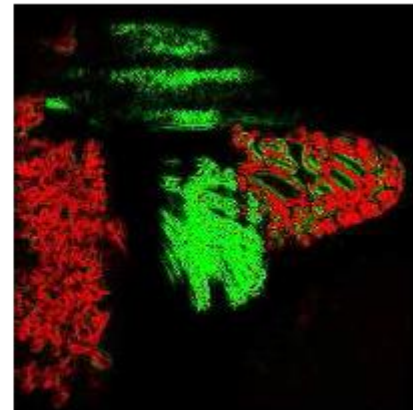
- Metryka mierzy zmiany jasności w sąsiedztwie każdego piksela
- Interpretacja wartości metryki:
 - Większa wartość oznacza mniejszy stopień rozmycia (a więc i lepszą jakość)



oryginalny



zdegradowany



Blurring Metric

Metryki bezwzględne

- **Blocking Metric**

- Metryka opisuje zblokowanie pikseli
- Zniekształcenie to wynika ze sposobu przetwarzania obrazu przy kompresji z wykorzystaniem transformaty DCT – odbywa się ono w blokach pikseli, zwykle szerokość i wysokość bloku jest wielokrotnością liczby 8
- Zniekształcenia tego typu są widoczne przy dużym stopniu kompresji, szczególnie na powierzchniach o zbliżonej (lecz nie identycznej barwie i jasności)
- Metryka mierzy różnice barwy i kontrastów pikseli położonych na brzegu sąsiednich bloków o rozmiarze 8 na 8 pikseli
- Wartość 0 oznacza brak zniekształceń – im większa wartość tym większe zniekształcenia (a niższa jakość, odwrotnie niż dla Blurring Metric)

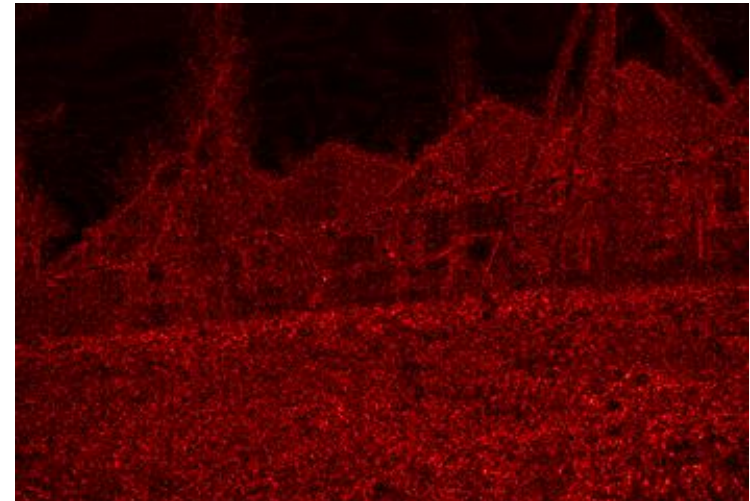
Metryki bezwzględne



Obraz oryginalny



Obraz zdegradowany



Mapa różnic

Metryki bezwzględne

- **Blocking Metric**



A



B

A – zakodowana ramka

B – wizualizacja metryki

Metryki bezwzględne

- **Brightness Flicking Metric**
 - Metryka opisuje zniekształcenia charakteryzujące się migotaniem obrazu (flicking) pomiędzy sąsiednimi ramkami
 - Wartość oznacza średnią miarę bezwzględnej różnicy jasności wszystkich pikseli obrazu, w porównaniu do poprzedniej ramki
 - Wyższa wartość oznacza większe migotanie

Metryki bezwzględne

- **Drop Frame Metric**

- Oblicza liczbę uszkodzonych ramek w sekwencji wideo
- Maksima dla ramek powtórzonych
- Oryginalna ramka i wizualizacja (kolor szary oznacza, że jasność piksela jest taka sama jak w poprzedniej ramce):



Dziękuję za uwagę.