

# RIG Z1.CD

## Cechy diagnostyczne metod analizy i przetwarzania obrazu stereoskopowego

© 2015 FINN Sp. z o.o. Wszelkie prawa zastrzeżone

### Historia zmian dokumentu:

Data	Wersja	Osoba	Opis
2014.04.18	1.0	Przemysław Sztoch Adam Wojciechowski Bartosz Sakowicz Przemysław Sękalski Dariusz Makowski Piotr Mazur Robert Ritter Andrzej Napieralski	Utworzenie struktury dokumentu.
2014.09.26	1.1	Agnieszka Rytel	Aktualizacja i weryfikacja wynikająca z przeprowadzonego przeglądu jakości
2014.12.03	1.2	Przemysław Sztoch Ryszard Sztoch Jakub Wawrzyniak	Aktualizacja i weryfikacja wynikająca z przeprowadzonego przeglądu jakości
2015.03.30	1.3	Ryszard Sztoch	Weryfikacja formalna dokumentu – nagłówki i stopki
2015.04.30	1.4	Adam Wojciechowski	Aktualizacja wynikająca z przeprowadzonych badań
2015.09.05	2.0	Marek Kamiński Rafał Kotas Paweł Marciniak Bartosz Sakowicz	Aktualizacja wynikająca z przeprowadzonych badań

1	Wstęp .....	4
1.1	Rejestracja obrazu stereoskopowego.....	4
1.2	Powiązanie z projektem ROS3D.....	4
2	Cechy diagnostyczne obrazu stereoskopowego .....	6
2.1	Kalibracja torów optycznych.....	6
2.2	Dysparycja pionowa.....	7
2.3	Dysparycja pozioma (paralaksa pozioma) .....	7
2.4	Rozbieżność pomiędzy akomodacją wzroku a konwergencją obrazów .....	7
2.5	Rozbieżności pomiędzy konwergencją a głębią ostrości kamer .....	7
2.6	Błędy w jasności/kontraście .....	8
2.7	Błędy w ostrości.....	8
2.8	Błędy w zakresie tonalnym obrazu .....	8
2.9	Błędy w budzecie paralaksy.....	8
2.10	Korekta trapezowości obiektów obrazu (ang. <i>keystoneeffect</i> ) .....	8
2.11	Szereg bezwzględnych miar jakości obrazu.....	8
2.12	Prędkość zmian obrazu pomiędzy klatkami.....	8
2.13	Obecność obiektów wystających z kadru (szczególnie pierwszoplanowych).....	8
3	Miary stosowane do oceny jakości obrazów 2D .....	10
3.1	PSNR ( <i>Peak signal-to-noise ratio</i> ) [1].....	10
3.2	SSIM ( <i>Single scale structural similarity</i> ) [2] .....	11
3.3	MSSIM ( <i>Multi-scale structural similarity</i> ) [3].....	12
3.4	VSNR ( <i>Visual signal-to-noise ratio</i> ) [4].....	12
3.5	VIF ( <i>Visual information fidelity</i> ) [5].....	14
3.6	UQI ( <i>Universal quality index</i> ) [6].....	14
3.7	IFC ( <i>Information fidelity criterion</i> ) [7].....	14
3.8	NQM ( <i>Noise quality measure</i> ) [8].....	15
3.9	WSNR ( <i>Weighted signal-to-noise ratio</i> ) [9] .....	15
3.10	PHVS ( <i>Modified PSNR based on HVS</i> ) [10].....	15
3.11	JND ( <i>Just noticeable distortion model</i> ) [11].....	15
3.12	Bibliografia.....	16
4	Miary stosowane do oceny jakości obrazów 3D .....	17
4.1	<i>Research on subjective stereoscopic image quality assessment</i> .....	17
4.2	<i>CYCLOP: A Stereo Color Image Quality Assessment Metric</i> .....	17
4.3	<i>Perceptual quality assessment for stereoscopic images based on 2D image quality metrics and disparity analysis</i> .....	18
4.4	<i>Towards compound stereo-video quality metric: a specific encoder-based framework</i> .....	18

4.5	<i>Objective quality assessment of depth image based rendering in 3DTV system</i> .....	18
4.6	<i>Video quality assessment: from 2D to 3D – challenges and future trends</i> .....	19
4.7	<i>Perceptual video quality metric for 3d video quality assessment</i> .....	19
4.8	<i>Stereoscopic images quality assessment</i> .....	19
4.9	<i>Quality of Experience Assessment for Stereoscopic Images</i> .....	19
4.10	<i>A comprehensive database and subjective evaluation methodology for quality of experience in stereoscopic video</i> .....	20
4.11	<i>Algorithmic assessment of 3D quality of experience for images and videos</i> .....	20
4.12	<i>Stereoscopic video quality assessment model based on spatial-temporal structural information</i> 20	
4.13	<i>Quality Assessment of Stereoscopic Images</i> .....	20
4.14	<i>Objective No-Reference Stereoscopic Image Quality Prediction, Based on 2D Image Features and Relative Disparity</i> .....	20
4.15	<i>3D-DCTbased perceptual quality assessment of stereo video</i> .....	20
4.16	<i>Paired comparison-based subjective quality assessment of stereoscopic images</i> .....	21

# 1 Wstęp

## 1.1 Rejestracja obrazu stereoskopowego

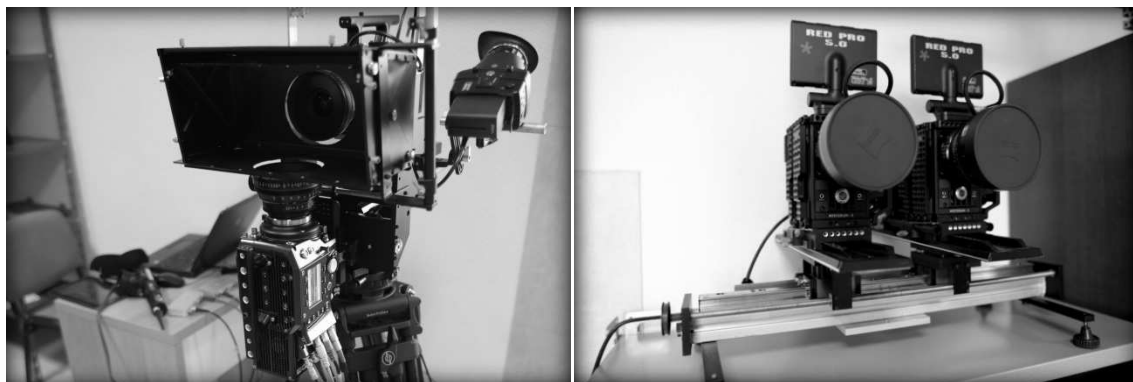
Rejestracja obrazu stereoskopowego 3D to zapisanie tego co widzimy naszym **dwójgiem** oczu.

Stereoskopia to widzenie przestrzenne (trójwymiarowe) polegające na analizie dwóch obrazów widzianych z perspektywy lewego i prawego oka. 3D to przestrzeń trójwymiarowa (wysokość, szerokość i głębokość).

Oczy są oddalone od siebie o około 6,5 cm. To jest powodem tego, że widzą tą samą przestrzeń z nieco innej perspektywy i z innego kąta. Mózg następnie łączy oba obrazy w jeden obraz przestrzenny. Nazywamy to widzeniem obuocznym.

Kiedy rejestrujemy obraz stereoskopowy określonych obiektów w przestrzeni 3D wykonujemy to przy pomocy dwóch kamer z rozsuniętymi wzajemnie osiami torów optycznych.

Gdy odtwarzamy obraz stereoskopowy musimy zapewnić, aby obraz zarejestrowany lewą kamerą trafił tylko do lewego oka, natomiast prawą kamerą był widziany tylko przez prawe oko. Mózg widza połączy oba płaskie obrazy dając poczucie głębi.



## 1.2 Powiązanie z projektem ROS3D

Niniejszy dokument sporządzono na podstawie dokumentacji opracowanej w trakcie realizacji projektu pt. „Rejestracja obrazu stereoskopowego” (w skrócie ROS3D), realizowanego w ramach Przedsięwzięcia pilotażowego Wsparcie badań naukowych i prac rozwojowych w skali demonstracyjnej „DEMONSTRATOR+”.

W projekcie wykonano prototypy 4 różnych konstrukcji rigów Systemu ROS3D, które poddano testowaniu podczas rejestracji materiału przede wszystkim o klatkarzu 25 i 50/sek. (25 i 50 fps), w niektórych przypadkach scen osiągającym wartość 100/sek. Wykonane 2 prototypy obejmowały rigi równoległe, natomiast 2 kolejne rozwinęły rigi lustrzane, zarówno w zakresie konstrukcji jak również w przedmiocie elektronicznych urządzeń (kontrolerów) oraz oprogramowania do analizy i konwersji obrazu 3D.

Motywacja do weryfikacji wyników badań uzyskanych w projekcie ROS3D wynikała z dwóch głównych przesłanek zdeterminowanych przez założone przeznaczenie riga, opisanych poniżej:

1. charakter rejestrowanych scen i parametry techniczne obrazu,
2. wynikające z pierwszej przesłanki odmienności konstrukcyjne riga oraz specjalizowane w kontekście rejestrowanych scen (wymaganych parametrów obrazu) kamery i obiektywy.

Niniejszy Projekt rozwijał konstrukcję riga lustrzanego mającą przeznaczenie do rejestracji obrazów stereoskopowych o innym charakterze od założonych w projekcie ROS3D. Mianowicie przedmiotowy dla niniejszego Projektu rig ultralekki (RU) umożliwia rejestrację obrazu o bardzo dynamicznej akcji

(uzyskiwanej dynamicznym sposobem prowadzenia kamer, a praktycznie riga z zainstalowanymi kamerami), w klatkarzu do 300 fps, m. in. przy pomocy zewnętrznych konstrukcji typu steadycam, easyrig, czy BotDolly wymagających ograniczenia wagi zestawu riga i pozostałego sprzętu do filmowania do około 20 kg.

Analizując rezultaty Projektu ROS3D, a w szczególności prototypową konstrukcję ROS.RL2 (Rig lustrzany mobilny – mały) postanowiono, że w niniejszym Projekcie rozwijana będzie konstrukcja riga metalowego oparta na innych założeniach wykonania mechanizmów ruchowych obsługujących zmianę wartości bazy i konwergencji. Mianowicie jedna z kamer została na stałe mocowana do konstrukcji riga, natomiast swoje położenia będzie zmieniać wyłącznie druga kamera. Uzasadnieniem takiej decyzji był wspomniany wyżej charakter zdjęć zaplanowanych do realizacji przy pomocy riga RU. Będą to plany bliskie, a bardzo rzadko średnie. Założono, że brak symetrii ustawienia kamer w stosunku do umownej osi rejestracji obrazu będzie bez znaczenia dla jakości rejestrowanego obrazu 3D. W pierwszej fazie badań i prac rozwojowych potwierdziły to analizy geometrii, a następnie wstępne próby na adaptowanej w tym celu konstrukcji riga RL2.

Zyskiem tak zmienionej ideowo konstrukcji riga okazało się przede wszystkim uproszczenie konstrukcji, zmniejszenie wagi, zwiększenie sztywności konstrukcji oraz zmniejszenie podatności na odkształcenia podczas dynamicznego używania całego zestawu do filmowania.

Zaplanowane i zrealizowane w niniejszym Projekcie działania dla uzyskania technologii analizy i przetwarzania obrazu stereoskopowego skutecznej w przypadku klatkarzu osiągającego wartość 300/sek. oraz realizowanej na bardzo wymagających scenach w zastosowaniach do filmów dla symulatorów i kin 5D/6D/7D oraz gier komputerowych, wykonanych metodą stereoskopową 3D w rozdzielczości UltraHD/4K, napotkały bariery innego rodzaju i o wyższej skali trudności niż występujące w Projekcie ROS3D.

Okazało się, że barierą obiektywną i niezależną do działań Projektu są parametry techniczne profesjonalnego sprzętu filmowego (kamer i obiektywów), a w szczególności jego waga i gabaryty oraz możliwe do uzyskania parametry techniczne rejestrowanego obrazu (rozdzielczość, klatkarz, zgodność wszystkich parametrów na dwóch kamerach/obiektywach z uwagi na efekt 3D).

Na przykład postanowiono, że Rig ultralekki będzie przystosowany do współpracy z kamerami RED Epic/Dragon. Waga całego zestawu do filmowania w tzw. minimalnej konfiguracji powinna nie przekraczać 18,8 kg przy zastosowaniu stosunkowo małych i lekkich obiektywów stałogniskowych. Okazało się nie możliwe wykonanie lekkiej i jednocześnie odpowiednio sztywnej konstrukcji riga do współpracy z kamerami PHANTOM Flex4K, których 2 szt. ważą 12,5 kg. Waga całego zestawu do filmowania przekracza 31 kg. Przyjęcie różnych tego typu założeń (ustaleń, zaleceń) dla sprzętu determinowało rozwijanie właściwych procedur rejestracji obrazu, w tym innego podejścia do szczegółowego planu czynności analizy i konwersji obrazu 3D, od realizowanych w projekcie ROS3D. W szczególności ciężar ustaleń konfiguracji riga i parametrów pozostałego sprzętu filmowego przesunięto na zdjęcia próbne, ponieważ z uwagi na wagę i wielkość zrezygnowano z instalacji na rigu ultralekkim elektronicznych urządzeń sterujących podczas zdjęć zasadniczych (na rzeczywistym planie zdjęciowym).

Przeprowadzone badania przemysłowe w niniejszym Projekcie w zadaniu nr 1 oraz prace rozwojowe wykonane w zadaniu nr 2 i 3 wykazały, że podwyższony klatkarz oraz inne bardziej wymagające warunki rejestracji obrazu nie zmniejszają wartości i zasadności wskazanych w projekcie ROS3D cech diagnostycznych metod analizy i przetwarzania obrazu stereoskopowego.

## 2 Cechy diagnostyczne obrazu stereoskopowego

### 2.1 Kalibracja torów optycznych

Tworzenie poprawnego obrazu stereoskopowego jest zagadnieniem, które wymaga uwzględnienia szeregu czynności zarówno na etapie przygotowania (preprodukcja), jak również na etapie realizacji (produkcja) i postprodukcji. Problemy w jakości obrazu mogą leżeć zarówno w dopasowaniu parametrów kamer jak i w samej jakości obrazu (kadr, szybkość prowadzenia kamer). Dopasowanie parametrów kamer realizowane jest zarówno na etapie wstępnym (kalibracja) jak i podczas realizacji filmowej. O ile na etapie przygotowania obsługa ma stosunkowo dużo czasu na skalibrowanie układu kamer, to na etapie realizacji tego czasu jest relatywnie mało.

Istotnym elementem na etapie przygotowania jest kalibracja ustawień riga.



Rozbieżności niektórych parametrów dla obrazów lewego i prawego oka jest bardzo trudno wyeliminować lub jest to niemożliwe.

Dotyczy to w szczególności riga lustrzanego, w którym kamery zazwyczaj są ustawione prostopadle. Wynika to z tego, że jeden obraz (zazwyczaj oka lewego) jest rejestrowany jako światło przepuszczone przez półprzepuszczalne lustro (zazwyczaj przez kamerę poziomą). Drugi obraz (oko prawe) jest rejestrowany jako światło odbite od półprzepuszczalnego lustra (zazwyczaj przez kamerę pionową). Wpływa to na różnice jasności, różnice w zakresie tonalnym oraz różne zabarwienie obu obrazów.

W przypadku rigu równoległego różnice mogą wynikać z różnego ustawienia kamer względem obecnych w scenie źródeł światła. Ponieważ rig równoległy jest zazwyczaj używany na otwartych przestrzeniach słońce jest bezpośrednią przyczyną powstawania refleksów i rozbieżności.

W trakcie wykonywania zdjęć na planie filmowym cały czas istnieje potrzeba korygowania nastaw kamer i riga, aby odpowiadały one aktualnie nagrywanym ujęciom. Wykonywane na bieżąco zmiany nastaw (np. baza stereoskopowa, konwergencja, ogniskowa, przysłona, ostrość, klatkaż nagrywania, czas ekspozycji) powinny ze sobą współgrać i spełniać odpowiednią korelację. Nieodpowiednie ustawienie bazy stereoskopowej względem planu zdjęciowego może skutkować niewłaściwą dysparcją obrazów (paralaksą) podczas projekcji, co będzie powodowało znaczące obniżenie komfortu odbioru obrazu stereoskopowego. Podobnie częste przestawianie parametrów riga oraz jego przenoszenie może powodować utratę kalibracji wertykalnej skutkującej dysparcją pionową, która również nie jest pożądana.

Niewłaściwe rozbieżności mogą pojawić się na polu dopasowania głębi ostrości obiektywów, która powinna zawierać się w głębi obrazu stereoskopowego.

Ewentualne błędy mogą pojawić się w obszarze rozbieżności akcji na planie filmowym, a obszarem koncentracji układu kamer. W tym aspekcie mamy, bowiem do czynienia z sytuacją, kiedy to reżyser/operator planuje i realizuje określony scenariusz chcąc pokazać widzowi pewną scenę. Nie

powinna pojawić się wtedy rozbieżność pomiędzy punktem kadru, w którym rozgrywa się akcja a punktem kadru, w którym koncentrują się kamery (mają ustawioną: linię optyczną, głębię ostrości, głębię obrazu).

Błędy w percepcji obrazu stereoskopowego mogą pojawić się również, gdy szybkość zmian na ekranie będzie ponadnaturalna, tak więc prędkość zmian położenia/orientacji rigu w rejestracji obrazu stereoskopowego powinna być niższa niż przy rejestracji obrazu 2D.

Cześć z błędów w obrazie stereoskopowym może powstać również na etapie postprodukcji. Należą do nich głównie nieskoordynowane zmiany jasności, kontrastu, zakresu tonalnego, dysparycji poziomej, pionowej itp. Jednakże większość z tych błędów jest odwracalna, w ramach historii projektu, pod warunkiem świadomości ich popełnienia.

## 2.2 Dysparycja pionowa

Pionowa dysparycja (przesunięcie) obrazów powinna być całkowicie zlikwidowana. W rzeczywistości odpowiada to sytuacji, w której widz ma jedno oko wyżej, a drugie niżej. Błąd ten może zostać usunięty w postprodukcji i może przykładowo wynikać ze złego skalibrowania kamer, niewłaściwego nałożenia obrazów w postprodukcji lub nawet niestarannego spasowania dwóch rzutników podczas projekcji w kinie. Błąd ten dość prosto daje się skorygować w postprodukcji. Nie mniej powinniśmy dołożyć niezbędnej staranności aby wyeliminować ten błąd już w trakcie samej rejestracji (na etapie produkcji) – z pewnością będzie to tańsze i szybsze.

## 2.3 Dysparycja pozioma (paralaksa pozioma)

Zbyt duża pozioma dysparycja (przesunięcie większe niż ok. +6,5 cm) powoduje, że oczy widza (kierunek patrzenia każdego oka) oglądającego materiał ustawiają się rozbieżnie. Wpływa to na złe samopoczucie, nudności i zawroty głowy. Aby zapewnić prawidłową zbieżność oczu podczas oglądania należy podczas rejestracji obrazu zapewniać odpowiedni budżet paralaksy obrazu i na etapie postprodukcji odpowiedni podział pomiędzy paralaksę ujemną (bliską) i dodatnią (daleką).

## 2.4 Rozbieżność pomiędzy akomodacją wzroku a konwergencją obrazów

Podczas oglądania filmu widz kieruje się emocjami i koncentruje wzrok na elementach, które z jego punktu widzenia (wynikającym z fabuły) są najistotniejsze. Najczęściej jest to zgodne z wizją reżysera/operatora, który koncentruje kadr na najistotniejszych, z punktu widzenia fabuły, elementach. W przypadku obrazu 3D należy dodatkowo zadbać, aby obiekty istotne z punktu widzenia fabuły (jak również przypuszczalnie widza) znajdowały się na płaszczyźnie zerowej paralaksy, gdzie dzięki konwergencji krzyżują się osie patrzenia kamer (w centrum ostrości obrazu). W skrajnym przykładzie przy rozbieżności akomodacji i konwergencji widz koncentruje się na „nieostrym” fragmencie obrazu, zaś ostrość jest ustawiona na mniej istotnych, z punktu widzenia fabuły, elementach. Jeśli plan będący przez dłuższy czas w sferze zainteresowań widza będzie znajdował się poza płaszczyzną zerowej paralaksy (przed lub w głębi sceny) to wtedy akomodacja wzroku, która koncentruje się na płaszczyźnie wyświetlania nie będzie skorelowana z konwergencją. Zjawisko to potocznie znane jest pod nazwą konfliktu akomodacji i konwergencji (ang. *vergence - accommodation conflict*).

## 2.5 Rozbieżności pomiędzy konwergencją a głębią ostrości kamer

Ustawiając lub planując konwergencję kamer w Rigu 3D niejako zakładamy, w jakim zakresie odległości od kamer obiekty planu będą mieściły się w budżecie paralaksy – będą widziane maksymalnie z założonym stopniem rozmycia (*patrz* krążek rozmycia). Wartość ta nakłada się na ustawienia samej kamery (ogniskowa, przysłona, itp.), które skutkują określoną głębią ostrości każdego obrazu z osobna. Nie dość, że głębia ostrości każdej z kamer powinna być taka sama to jeszcze powinna ona pokrywać się ostrością stereoskopową wynikającą z konwergencji obrazów.

Projekt pt. *Osprzęt i oprogramowanie automatyzujące proces filmowania obrazu 3D* nr Umowy z NCBR: *INNOTECH-K3/H13/16/227587/NCBR/14* dofinansowany w ramach Programu „*INNOTECH*” w ścieżce programowej *HI-TECH*, współfinansowany ze środków Narodowego Centrum Badań i Rozwoju oraz Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego.

## 2.6 Błędy w jasności/kontraście

Obrazy pochodzące z dwóch kamer powinny mieć odpowiednio identyczną (z pewną tolerancją) ogólną jasność oraz kontrast obrazu. Różnica w jasności lub kontraście może wynikać przykładowo z ustawienia źródeł światła na planie lub samej konstrukcji riga (np. jeden obraz jest rejestrowany przez półprzepuszczalne lustro, a drugi jako obraz odbity od tego lustra).

## 2.7 Błędy w ostrości

Jednym z błędów ostrości, poza różnymi ustawieniami kamer może być koncentracja ostrości symetryczna bądź asymetryczna (nie pokrywająca się w płaszczyźnie obrazu), jak również sama ostrość i jej głębia. Ten błąd w niewielkim zakresie daje się skorygować w postprodukcji.

## 2.8 Błędy w zakresie tonalnym obrazu

Jeśli z niewyjaśnionych przyczyn zakres tonalny obrazów pochodzących z dwóch kamer się różni to ciężko jest skorygować te błędy, chyba że znamy/znajdziemy korekcję barwną ujednociającą obrazy.

## 2.9 Błędy w budżecie paralaksy

Jeśli budżet paralaksy jest zbyt duży to bardzo trudno jest wyeliminować błędy paralaksy za pomocą samej zmiany proporcji pomiędzy paralaksa ujemną a paralaksą dodatnią.

## 2.10 Korekta trapezowatości obiektów obrazu (ang. *keystoneeffect*)

Trapezowatość obrazu obiektów sceny jest zjawiskiem, które pojawia się w sytuacji, gdy określony obiekt rejestrowany jest dwoma kamerami z dwóch różnych perspektyw (pozycji). Czym większa względna odległość torów optycznych tym efekt trapezowatości staje się mocniej dokuczliwy, szczególnie w pierwszym planie sceny. Efekt można niwelować zachowując odpowiednią relację bazy rigu i odległości od pierwszego planu. Jest również możliwość geometrycznego skorygowania i ujednoczenia perspektywy dwóch obrazów obiektu (lewego i prawego), jednak wymaga to przekształcenia geometrycznego obrazów, a następnie ich kadrowania.

## 2.11 Szereg bezwzględnych miar jakości obrazu

Istnieje szereg miar jakości obrazu 2D, z których niektóre okazują się jako mające wysoką korelację z subiektywnymi odczuciami widza. Należą do nich VQM (*Video Quality Metric*) [1], SSIM, UQI, C4, RRIQA [2], *binocular distortions* [3,4,5,6], itp.

## 2.12 Prędkość zmian obrazu pomiędzy klatkami

Parametr ten ma znaczenie z punktu widzenia percepcji widza. Punktem odniesienia jest szybkość obserwowalnych przez człowieka zmian zachodzących w naturze, która pomimo teoretycznie ograniczonych prędkością światła wartości jest znacznie niższa. Istnieje dla każdej osoby pewien subiektywny próg, powyżej którego zmiany zachodzą za szybko i percepcja (zrozumienie) zachodzących zmian jest ograniczone. Bazując na opinii ekspertów/badaniach na osobach można pokusić się o wskazanie granicznych wartości progów zmian zachodzących dla poszczególnych planów, powyżej których obraz zaczyna być „drażniący”.

## 2.13 Obecność obiektów wystających z kadru (szczególnie pierwszoplanowych)

W sytuacji, gdy pierwszoplanowy obiekt wystaje z kadru (jest częściowo widoczny) obserwator ma problem z pełnym określeniem głębi danego obiektu. Najczęściej głębia obiektu w scenie kłóci się z głębią fizyczną ekranu, na którym oglądany jest film. Poza tym, jeśli obiekt wchodzi w kadr tylko jednej kamery, to znaczy, że będzie widoczny tylko przez jedno oko obserwatora i wtedy będzie mu się on wydawał przezroczysty. Ujęcia z takimi klatkami powinny być skorygowane lub odpowiednio skadrowane.



Jeżeli obraz rejestrujemy kamerami ustawionymi równolegle (brak konwergencji) to musimy być świadomi o konieczności skorygowania konwergencji w postprodukcji. Będzie to wiązało się z koniecznością dodatkowego przycięcia obrazów, co zwiększa ryzyko, że obiekty z bliskiego planu będą widziane tylko „jednym okiem”. Rejestracja za pomocą rigu równoległego wymaga zatem założenia odpowiedniego zapasu (marginesu) w kadrze.

## 3 Miary stosowane do oceny jakości obrazów 2D

Obiektywne miary jakości obrazu można podzielić na dwie klasy w oparciu o sposób ich obliczania.

Pierwszą klasę stanowią metody oparte o łatwą w obliczeniach formułę matematyczną. Należą do nich powszechnie używane:

- średni błąd kwadratowy (MSE),
- stosunek szczytu sygnału do szumu (PSNR),
- pierwiastek błędu średniokwadratowego (RMSE),
- błąd bezwzględny (MAE),
- stosunek sygnału do szumu (SNR).

Druga klasa metod opiera się na modelu ludzkiego wzroku (HVS) i włącza do oceny cechy związane z percepcją. Niestety, nie wykazano w literaturze by któraś ze skomplikowanych metod tej klasy miała istotną przewagę nad prostymi metodami pierwszej grupy w dowolnych, określonych warunkach testowych i przy różnych źródłach zniekształceń obrazu. Proste metryki pierwszej grupy są nadal atrakcyjne z dwóch powodów. Po pierwsze, są one łatwe do obliczenia i najczęściej mają niską złożoność obliczeniową. Po drugie, są one niezależne od indywidualnych cech obserwatora i warunków prowadzenia pokazu. Uważa się, że warunki obserwacji odgrywają ważną rolę w ludzkiej percepcji jakości obrazu, niemniej jednak są w większości przypadków niedostępne dla systemu analizy. Jeśli dane takie są dostępne wyniki analiz są kłopotliwe w użyciu i zmuszają użytkownika do wykonywania licznych pomiarów.

### 3.1 PSNR (*Peak signal-to-noise ratio*) [1]

PSNR jest wciąż jedną z najczęściej stosowanych metryk oceny jakości wizyjnej obrazu i jest nadal powszechnie używany do oceny skuteczności kompresji i transmisji sygnałów. Chociaż wydajność PSNR jest gorsza niż w wielu innych metryk (zwłaszcza w kontekście niektórych rodzajów zniekształceń), jest on nadal atrakcyjny ze względu na dużą łatwość obliczeń i wynikającą z tego możliwość szerszego użycia np. w optymalizacji.

PSNR najłatwiej jest zdefiniować przy wykorzystaniu błędu średniokwadratowego (MSE). Biorąc pod uwagę, wolny od szumów monochromatyczny obraz  $X$  o wymiarach  $M \times N$  i jego zniekształconą wersję  $X'$ , MSE i PSNR jest zdefiniowany jako:

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (x_{ij} - x'_{ij})^2$$

$$PSNR = 10 \log_{10} \frac{I^2}{MSE} [db]$$

$$I = 2^B - 1$$

$B$  – liczba bitów na pixel

$M, N$  – wymiary obrazu

$x_{ij}$  – piksel  $(i, j)$  oryginalnego obrazu

$x'_{ij}$  – piksel  $(i, j)$  zakłóconego obrazu

PSNR jest najczęściej używany do pomiaru jakości rekonstrukcji kodeków kompresji stratnej. Sygnałem w tym przypadku są oryginalne dane, a szum to błąd wprowadzony przez kompresję. Wyższa wartość PSNR ogólnie wskazuje, na wyższą jakość rekonstrukcji. W niektórych przypadkach wskaźnik ten może być jednak

bardzo mylący i jego zastosowanie wymaga dużej ostrożności. Stosując różne techniki przetwarzania danych do obrazu wyjściowego można wprowadzić różnego rodzaju zakłócenia i dobierając odpowiednio parametry przetwarzania łatwo uzyskać obrazy o takim samym współczynniku PSNR, znacznie jednak różniące się w odbiorze subiektywnym. Zgodnie z zaleceniami PSNR może być stosowany do porównywania operacji przetwarzania danych opartych na zbliżonych algorytmach.

Głównym ograniczeniem PSNR jest to, że opiera się ona wyłącznie na porównaniu numerycznym i nie uwzględnia żadnego poziomu czynników biologicznych ludzkiego wzroku, takich jak np. indeks podobieństwa strukturalnego (SSIM). W przypadku obrazów barwnych, MSE jest obliczane dla poszczególnych kanałów i uśrednione po ich liczbie. Innym rozwiązaniem może być wykonanie PSNR na obrazie w skali odcieni szarości, zwłaszcza po uwzględnieniu faktu lepszej czułości oka dla tego typu obrazów.

Typowe wartości PSNR w stratnej kompresji obrazu i wideo mieszczą się w przedziale 30 - 50 dB przy głębi bitowej 8 bitów. Dla danych 16 bitowych typowe wartości PSNR mieszczą się pomiędzy 60 a 80 dB. Za dopuszczalną wartość utraty jakości transmisji bezprzewodowej jest uważany przedział około 20 dB do 25 dB. Podczas porównywania dwóch identycznych obrazów MSE będzie wynosić zero, a więc PSNR będzie wtedy niezdefiniowany (dzielenie przez zero).

W przypadku sekwencji wizyjnych PSNR podaje się niezależnie dla każdego obrazu (np. w formie wykresu), lub wyznacza uśrednioną wartość dla wszystkich obrazów sekwencji.

### 3.2 SSIM (*Single scale structural similarity*) [2]

SSIM (indeks strukturalny) jest miarą porównującą informację strukturalną pomiędzy obrazem odniesienia i obrazem zniekształconym. Indeks opiera się na założeniu, że ludzki wzrok jest najlepiej przystosowany do wydobycia właśnie informacji strukturalnej (związków pomiędzy blisko położonymi względem siebie pikselami obrazu). Indeks uwzględnia informacje o jasności (średnie), kontraście (wariancje) i strukturze (kowariancje).

SSIM jest obliczana dla przesuwanego (piksel po pikselu) okna o wymiarach  $N \times N$  i środka w punkcie  $(x, y)$ . Zwykle stosuje się okno o wymiarach  $8 \times 8$ .

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

$\mu_x, \mu_y$  – wartości średnie

$\sigma_x, \sigma_y$  – wariancja

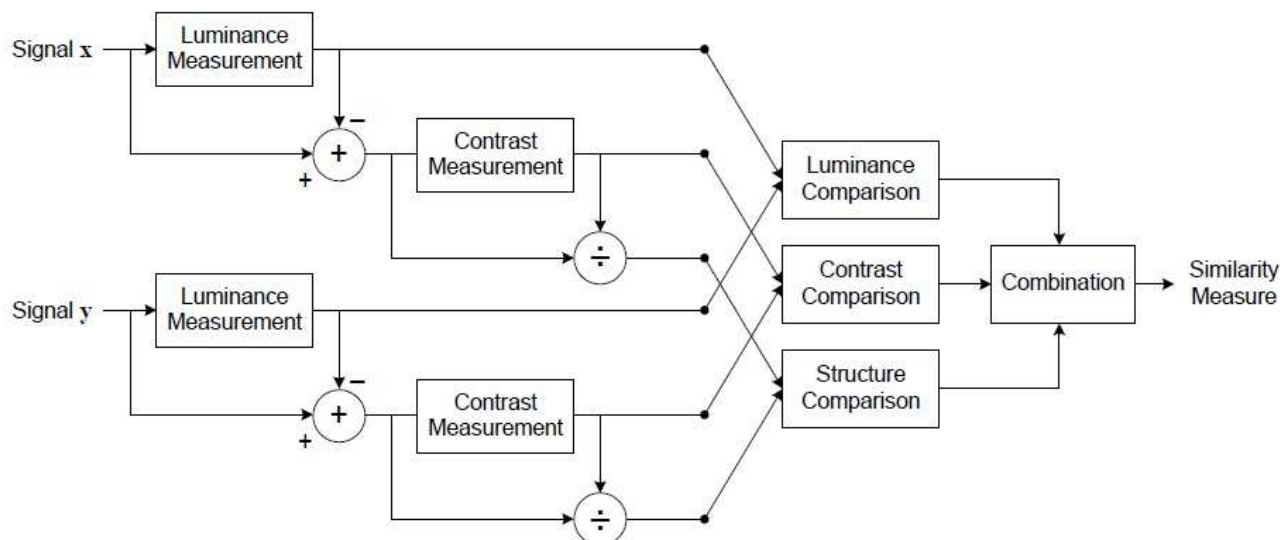
$\sigma_{xy}$  – kowariancja

$C_1, C_2$  – współczynniki stabilizujące (w celu uniknięcia zera w mianowniku)

Otrzymany wskaźnik SSIM jest wartością dziesiętną pomiędzy -1 i 1, a wartość 1 jest dostępna tylko w przypadku dwóch identycznych zestawów danych.

SSIM jest określany jako lepszy od tradycyjnych metod, takich jak szczyt sygnału do szumu (PSNR) czy błąd średniokwadratowy (MSE).

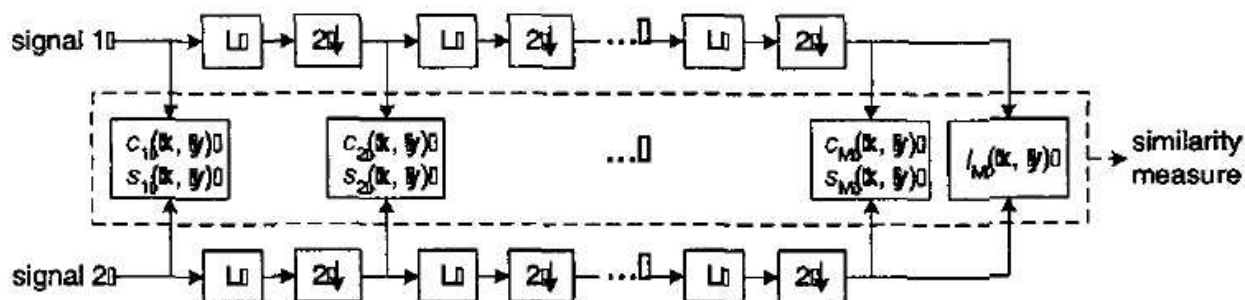
Koncepcyjny schemat blokowy SSIM:



### 3.3 MSSIM (Multi-scale structural similarity) [3]

MSSIM jest wskaźnikiem będącym rozszerzeniem SSIM. MSSIM iteracyjnie stosuje filtr dolnoprzepustowy w odniesieniu do oryginalnych i zmniejszonych obrazów. W każdej iteracji (skali obrazu) jest porównywany kontrast i struktura. Porównanie luminancji jest obliczane tylko dla ostatniej iteracji.

Schemat systemu przedstawiono na poniższym rysunku:



Ogólną ocenę uzyskuje się przez połączenie pomiarów przy różnych wartościach skali:

$$MSSIM(x, y) = |I_M(x, y)|^{\alpha_M} \prod_{j=1}^M |c_j(x, y)|^{\beta_M} |s_j(x, y)|^{\gamma_M}$$

$\alpha, \beta, \gamma$  - dobierane współczynniki

Odbiór szczegółów zdjęcia, zależy od rozlicznych parametrów takich jak gęstości próbkowania sygnału, odległości płaszczyzny obrazu do obserwatora i percepcyjnych możliwości obserwatora. W praktyce, subiektywna ocena danego obrazu zmienia się gdy czynniki te się różnią. Metoda pojedynczej skali może być poprawna tylko dla konkretnych ustawień. Metoda skali wielokrotnej jest wygodnym sposobem na włączenie szczegółów obrazu o różnej rozdzielczości.

### 3.4 VSNR (Visual signal-to-noise ratio) [4]

VSNR jest wskaźnikiem opartym na dekompozycji falkowej uwzględniającym model ludzkiego wzroku. Wymaga wstępnego przetworzenia danych, a następnie dwóch etapów obliczeniowych:

### 1) Przetwarzanie wstępne.

obliczyć zakłócenia  $E = I' - I$

$I$  – obraz oryginalny

$I'$  – obraz zakłócony

wykonać dyskretną transformatę falkową obrazów  $I$  i  $E$  w celu uzyskania zbioru podpasów dekompozycji falkowej.

obliczyć wektor częstotliwości przestrzennej:

$$f_m = 2^{-m} r v \tan\left(\frac{\pi}{180}\right)$$

$m = 1, 2, \dots, M$

$r$  – rozdzielczość (liczba pikseli na jednostkę długości)

$v$  – odległość obserwatora

### 2) Etap 1: Ocena wykrywalności zakłóceń

dla każdego  $f_m$  obliczyć próg detekcji kontrastu:

$$CT(E_f | I) = \frac{C(I_f)}{a_0 f^{a_2} \ln(f) + a_1}$$

$C$  – kontrast

dla każdego  $f_m$  zmierzyc rzeczywisty kontrast zniekształceń.

Jeżeli zniekształcenia są poniżej progu wykrywalności zniekształcony obraz jest uważany za doskonały ( $VSNR = \infty$ ), a algorytm zostaje zakończony.

### 3) Etap 2: Oblicz VSNR

obliczyć postrzegany kontrast zniekształceń  $d_{pc}$

obliczyć globalny kontrast zniekształceń  $d_{gp}$ .

obliczyć VSNR

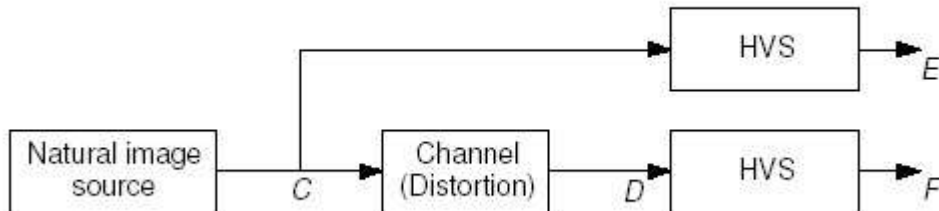
$$VNSR = 20 \log_{10} \left( \frac{C(I)}{\alpha d_{pc} + (1 - \alpha) \frac{d_{gp}}{\sqrt{2}}} \right)$$

VSNR ma stosunkowo niską złożoność obliczeniową i niskie wymagania dotyczące pamięci.

Tak jak większość algorytmów, które opierają się na modelach ludzkiego wzorku, VSNR ma istotne ograniczenia. Podstawową jego wadą jest to, że jest ograniczony tylko do obrazów w skali szarości. Innym ograniczeniem metryki VSNR jest brak lokalizacji przestrzennej zniekształceń. VSNR jest również wrażliwy na zniekształcenia geometryczne takie jak przesunięcie przestrzenne i obroty mające niewielki wpływ na wskaźniki wierności wizualnej.

### 3.5 VIF (Visual information fidelity) [5]

VIF (Wskaźnik wierności informacji wizualnej) jest obliczany na podstawie utraty informacji o obrazie z porównania obrazu oryginalnego i zniekształconego poddanego transformacji w modelu ludzkiego wzroku:



VIF oparty jest na ogół na dekompozycji falkowej. Może być też obliczany w dziedzinie płaszczyzny obrazu dla pikseli znajdujących się wewnątrz okna. Zamiast podpasm dekompozycji falkowej używane są wtedy współczynniki skali obrazu.

$$VIF = \frac{\sum I(C, F)}{\sum I(C, E)}$$

I – współczynnik wzajemnej wymiany informacji

VIF może być łatwo rozszerzony na wiele kanałów przy założeniu, że każdy kanał jest całkowicie niezależny od innych.

### 3.6 UQI (Universal quality index) [6]

UQI (Uniwersalny wskaźnik jakości) jest oparty na kombinacji trzech czynników: utraty korelacji, zniekształceniu luminancji i zniekształceniu kontrastu. Pod wieloma względami przypomina wskaźnik SSIM. Chociaż jest zdefiniowany według matematycznej formuły uwzględnia model ludzkiego wzroku.

Określenie „uniwersalny” oznacza że nie jest przeznaczony dla szczególnej klasy obrazów, a jego wartość nie zależy od warunków prezentacji i indywidualnych cech obserwatora. UQI zapewnia miarodajne wyniki przy różnych typach zakłóceń obrazu.

$$UQI = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2xy}{x^2 + y^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}$$

### 3.7 IFC (Information fidelity criterion) [7]

IFC (wskaźnik wierności informacji) jest oparty na bardzo podobnej koncepcji jak VIF.

$$IFC = \sum I(C, D)$$

Wartości IFC mieszczą się w przedziale od zera (brak wierności) do nieskończoności (doskonała wierność).

Do głównych zalet IFC należą brak parametrów związanych z modelem ludzkiego wzroku, czy parametrami urządzenia wyświetlającego (odległość widzenia, rozdzielczość ekranu). Czyni to z IFC metrykę łatwą w wyznaczeniu konkurencyjną do innych wskaźników typu FR.

### 3.8 NQM (Noise quality measure) [8]

NQM jest miarą opartą o ocenę jakości dodanego szumu. Bazuje na wykorzystaniu wariancji kontrastu, zmienności lokalnej jasności, przestrzenno-częstotliwościowym modelu percepcji i efekcie maskowania krawędzi.

$$NQM(dB) = 10 \log_{10} \left( \frac{\sum O_s^2(x, y)}{\sum (O_s(x, y) - I_s(x, y))^2} \right)$$

OS(x,y) – obraz zniekształcony

IS(x,y) – obraz wejściowy

### 3.9 WSNR (Weighted signal-to-noise ratio) [9]

WSNR (ważony stosunek sygnału do szumu) jest obliczany w dziedzinie częstotliwości. Różnica między obrazem odniesienia i obrazem zniekształconym jest przekształcana na dziedzinę częstotliwości z zastosowaniem dyskretnej transformaty Fouriera, a następnie mnożona przez funkcję czułości kontrastu.

$$WSNR(dB) = 10 \log_{10} \left( \frac{\sum |X(u, v) C(u, v)|^2}{\sum |(X(u, v) - Y(u, v)) C(u, v)|^2} \right)$$

X(u,v) – DFT sygnału oryginalnego

Y(u,v) – DFT sygnału zniekształconego

C(u,v) – DFT funkcji kontrastu

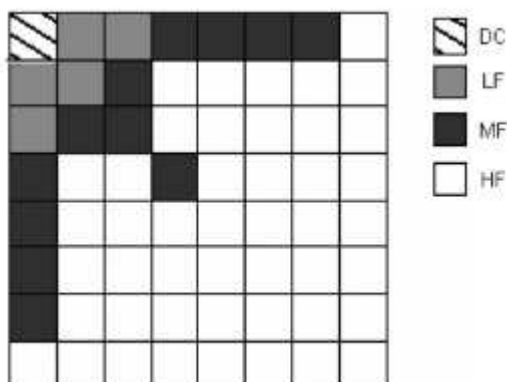
### 3.10 PHVS (Modified PSNR based on HVS) [10]

PHVS jest modyfikacją wskaźnika PSNR bazującą na modelu ludzkiego wzroku. Uwzględnienia model maskowania krawędzi wykorzystując dyskretną transformatę kosinusową.

### 3.11 JND (Just noticeable distortion model) [11]

JND jest wskaźnikiem bardzo silnie bazującym na modelu ludzkiego wzroku. Oparty jest na założeniu, że jedynie zauważalne zniekształcenia muszą być analizowane, a sygnały poza granicą ludzkiej percepcji mogą być pominięte. JND stanowi kombinację różnych czynników: czułości kontrastu, maskowania krawędzi (według przestrzenno-częstotliwościowego modelu percepcji opartego na transformacie kosinusowej), maskowania czasowego (związanego ze zmianami jasności).

$$JND(n, i, j, t) = T(n, i, j, t) \cdot a_{LUM}(n, t) \cdot a_{int ra}(n, i, j, t) \cdot a(n, t)$$



Powyższy rysunek przedstawia przykład modelu maskowania krawędzi (z podziałem na poszczególne składowe częstotliwościowe).

### 3.12 Bibliografia

- [1] Almohammad, Adel, and Gheorghita Ghinea. "Stego image quality and the reliability of PSNR." Image Processing Theory Tools and Applications (IPTA), 2010 2nd International Conference on. IEEE, 2010.
- [2] Pinson, M., Wolf, S.: A new standardized method for objectively measuring video quality. IEEE Trans. Broadcasting 50(33), 312–322 (2004)
- [3] Wang, Zhou, Eero P. Simoncelli, and Alan C. Bovik. "Multiscale structural similarity for image quality assessment." Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on. Vol. 2. IEEE, 2003.
- [4] Chandler, Damon M., and Sheila S. Hemami. "VSNR: A wavelet-based visual signal-to-noise ratio for natural images." Image Processing, IEEE Transactions on 16.9 (2007): 2284-2298.
- [5] Sheikh, Hamid R., and Alan C. Bovik. "A visual information fidelity approach to video quality assessment." The First International Workshop on Video Processing and Quality Metrics for Consumer Electronics. 2005.
- [6] Wang, Zhou, and Alan C. Bovik. "A universal image quality index." Signal Processing Letters, IEEE 9.3 (2002): 81-84.
- [7] Sheikh, Hamid R., Alan C. Bovik, and Gustavo De Veciana. "An information fidelity criterion for image quality assessment using natural scene statistics." Image Processing, IEEE Transactions on 14.12 (2005): 2117-2128.
- [8] Damera-Venkata, Niranjan, et al. "Image quality assessment based on a degradation model." Image Processing, IEEE Transactions on 9.4 (2000): 636-650.
- [9] Kim, Cheonshik. "Data hiding based on compressed dithering images." Advances in Intelligent Information and Database Systems. Springer Berlin Heidelberg, 2010. 89-98.
- [10] Ponomarenko, N., Battisti, F., Egiazarian, K., Carli, M., Astola, J., Lukin, V.: On between-coefficient contrast masking of DCT basis functions. In: CD-ROM Proc. VPQM, Scottsdale, Arizona, USA (2007)
- [11] Yang, X. K., et al. "Just noticeable distortion model and its applications in video coding." Signal Processing: Image Communication 20.7 (2005): 662-680.



## 4 Miary stosowane do oceny jakości obrazów 3D

Autorzy niniejszego raportu dokonali dogłębnego przeglądu literaturowego związanego z metodami oceny jakości obrazów 3D. Podsumowując wyniki studiów literaturowych, można sformułować następujące wnioski:

1. Miary stosowane do oceny obrazów 2D, bez wprowadzania dodatkowych współczynników i miar, nie nadają się do oceny obrazów 3D.
2. Zostało wykonanych wiele prób opracowania miar, które mogłyby być stosowane do oceny obrazu 3D, tym niemniej żadna z nich nie została przebadana na dostatecznie dużej próbie statystycznej, tak, aby przedstawiane wyniki mogły być uznane za wiarygodne.
3. W chwili powstawania niniejszego raportu nie istnieją uniwersalne miary umożliwiające obiektywną ocenę obrazów 3D.

Szczegółowy przegląd literaturowy zamieszczono poniżej.

### 4.1 *Research on subjective stereoscopic image quality assessment*

Xu Wang, Mei Yu, You Yang, and Gangyi Jiang, Faculty of Information Science and Engineering, Ningbo University, Ningbo, 315211, China

Artykuł przedstawia wstępne prace w ocenie jakości obrazów stereoskopowych. Autorzy koncentrują się na metryce subiektywnej powstałej w wyniku badań ankietowych na grupie 20 osób. Ocenie poddano zestaw 8000 obrazów do których sztucznie dodano wybrany typ zakłóceń.

W pracy zbadano cztery rodzaje zakłóceń:

- kompresje JPEG2000,
- kompresje JPEG,
- szum biały,
- rozmycie gaussowskie.

Opracowany wskaźnik subiektywny (DMOS) porównano ze stosowną miarą ilości dodawanych do obrazów zakłóceń (np. stopniem kompresji). W drugim etapie prac porównano DMOS ze wskaźnikiem PSNR stosowanym zazwyczaj do oceny obrazów dwuwymiarowych. W obu przypadkach wyniki były całkowicie zgodne z oczekiwaniami. Wyższy poziom zakłóceń to zarówno niższa wartość wskaźnika DMOS jak i PSNR. Zależność jest bliska liniowej. Wnioski takie są w zasadzie oczywiste i niewiele wnoszą do tej tematyki. Autorzy zapowiadają, że w ramach dalszych prac zajmą się już opracowaniem obiektywnego wskaźnika do oceny obrazów 3D.

### 4.2 *CYCLOP: A Stereo Color Image Quality Assessment Metric*

Aldo Maalouf, Mohamed-Chaker Larabi Geenov lab, Groupe Glaizer, 43 Rue Pierre Valette, 92240 Malakoff, France, aldo.maalouf@geenov.com, † XLIM Laboratory, UMR CNRS 6172, University of Poitiers, 2 Bd Marie et Pierre Curie, PO Box 30179, 86962 Futuroscope Chasseneuil, France, larabi@sic.univ-poitiers.fr

W publikacji autorzy po obszernym wstępie teoretycznym (odnoszącym się do typów metryk: *full reference*, *reduced reference* i *no reference*) wprowadzają nowy wskaźnik oceny obrazów 3D - SCIQA (*stereoscopic color image quality assessment*). Nowa miara zostaje poddana typowym testom porównawczym ze wskaźnikiem subiektywnym DMOS na wskazanej grupie zdjęć. Należy przyznać, że osiągnięte wyniki sugerują wysoką jakość proponowanego wskaźnika (wysoką korelację z DMOS). Należy jednak zaznaczyć, z czego zdają sobie sprawę również autorzy publikacji, że wskaźnik został sprawdzony

wyłącznie na grupie zdjęć. Prawdziwym wyzwaniem będzie dopiero przetestowanie go w sekwencji filmowej.

#### 4.3 *Perceptual quality assessment for stereoscopic images based on 2D image quality metrics and disparity analysis*

Junyong You<sup>1</sup>, Liyuan Xing<sup>1</sup>, Andrew Perkiś<sup>1</sup>, Xu Wang<sup>2</sup> <sup>1</sup>Centre for Quantifiable Quality of Service in Communication Systems (Q2S)\*, Norwegian University of Science and Technology (NTNU); <sup>2</sup>Ningbo University.

W publikacji autorzy przedstawili kompleksowe badania możliwości zastosowania miar do oceny obrazów 2D również w ocenie obrazów 3D. Przetestowane miary takie jak:

- *PSNR Peak signal-to-noise ratio*
- *SSIM Single scale structural similarity*
- *MSSIM Multi-scale structural similarity*
- *VSNR Visual signal-to-noise ratio*
- *VIF Visual information fidelity*
- *UQI Universal quality index*
- *IFC Information fidelity criterion*
- *NQM Noise quality measure*
- *WSNR Weighted signal-to-noise ratio*
- *PHVS Modified PSNR based on HVS*
- *JND Just noticeable distortion model*

nie nadają się do oceny obrazów 3D. Autorzy podjęli również próbę opracowania nowej, lepszej miary. W tym celu dokonali sprawdzenia wskaźników opartych na kombinacje wybranych wskaźników 2D z oceną różnicy obrazów prawego i lewego. Wyniki wydają się obiecujące jednak jak podkreślają sami autorzy jest to dopiero wstępna koncepcja i do opracowania efektywnego wskaźnika bardzo daleko.

#### 4.4 *Towards compound stereo-video quality metric: a specific encoder-based framework*

Atanas Boev<sup>1</sup>, Atanas Gotchev<sup>1</sup>, Karen Egiazarian<sup>1</sup>, Anil Aksay<sup>2</sup>, Gozde Bozdagi Akar<sup>2</sup> <sup>1</sup>Institute of Signal Processing, Tampere University of Technology, Tampere, Finland [firstname.lastname@tut.fi](mailto:firstname.lastname@tut.fi) <sup>2</sup>Electrical & Electronics Eng. Department, Middle East Technical University, Ankara, Turkey, [{anil,bozdagi}@eee.metu.edu.tr](mailto:{anil,bozdagi}@eee.metu.edu.tr)

Publikacja prowadzi do podobnych wniosków jak poprzednia. Miary oceny obrazów 2D nie nadają się bezpośrednio do oceny obrazów 3D. W publikacji dość szczegółowo wymieniono różnego rodzaju problemy, które sprawia ocena obrazów 3D nie występujące w obrazach 2D (np.: dyskomfort wizualny). Rozwiązanie problemu autorzy upatrują we wskaźniku będącym jakąś kombinacją wskaźników monoskopowych (QM) i stereoskopowych (QS). Jako wskaźnik QM autorzy najchętniej widzą SSIM, a jako QS różnicę obrazów prawego i lewego. Testy są wykonane na relatywnie bardzo małej grupie osób (16) i mimo obiecujących wyników nie dają definitywnych rozwiązań.

#### 4.5 *Objective quality assessment of depth image based rendering in 3DTV system*

Hang Shao, Xun Cao, Guihua Er Broadband Networks & Digital Media Laboratory TN List and Department of Automation, Tsinghua University 100084 Beijing, P.R. China

Publikacja po dość standardowym wstępie omawiającym problemy oceny obrazów 3D przedstawia autorski wskaźnik CSED (*Color and Sharpness of Edge Disortion*) bazujący na kombinacji składników oceniających zniekształcenie kolorów i zniekształcenie krawędzi. Autorzy oceniają wskaźnik jako wysoce efektywny co starają się pokazać w pracy poprzez standardowe porównanie z miarą subiektywną. Analogicznie jak w publikacji [2] ważnym składnikiem wskaźnika jest ocena barw, co stanowi pewne odwrócenie w stosunku do wskaźników 2D gdzie kolor był traktowany jako mniej użyteczny. Wskaźnik niewątpliwie wymaga dokładniejszego zbadania.

#### 4.6 Video quality assessment: from 2D to 3D – challenges and future trends

Quan Huynh-Thu Patrick Le Callet, Marcus Barkowsky Technicolor<sup>1</sup>, av. de Belle Fontaine CS 17616 – 35576 Cesson-Sévigné, France Polytech' Nantes IRCCyN UMR 6597 CNRS Rue Christian Pauc, 44306 Nantes, France

Publikacja [6] przedstawia całościowo problem oceny obrazów 3D. Wskazuje, że w stosunku do bogato opracowanej oceny obrazów 2D problem ten jest dopiero w fazie wstępnych badań i wymaga dalszych znaczących prac. Symptomatyczne jest to, że już nawet na etapie oceny subiektywnej pojawiają się wątpliwości. Obok wizualnych wrażeń obserwatora pojawia się parametr komfortu (rozumiany jako brak efektów typu ból głowy czy nudności), który również powinien zostać uwzględniony. Brak określonych wskaźników subiektywnych w oczywisty sposób utrudnia znalezienie wskaźników obiektywnych. Praca pozostawia większość zagadnień otwartych.

#### 4.7 Perceptual video quality metric for 3d video quality assessment

P. Joveluro, H. Malekmohamadi, W.A.C Fernando, and A.M. Kondoz, I-Lab Multimedia Communication Research, Faculty of Engineering and Physical Sciences, CVSP, University of Surrey, Guildford, UK. {P.Joveluro, H. Malekmohamadi, W.Fernando, A.Kondoz}@surrey.ac.uk

Praca pod względem stawianych celów i osiągniętych wyników bardzo przypomina publikację [3]. Stanowi próbę sprawdzenia zastosowania miar 2D (PQM i VQM) w problemie 3D. Wskaźniki porównywane są ze wskaźnikiem subiektywnym DMOS uzyskanym na bazie 30 ochotników. Podobnie jak [3] wyniki nie są zbyt obiecujące.

#### 4.8 Stereoscopic images quality assessment

Patrizio Campisi\*, Patrick Le Callet+, and Enrico Marini\*+ \*Dipartimento di Elettronica Applicata, Università degli Studi di Roma "Roma Tre", Via della Vasca Navale 84, 00146 Roma, Italy phone: +39.06.55177064, fax: +39.06.55177026, e-mail: campisi@uniroma3.it +IRCCyN, Ecole polytechnique de l'Université de Nantes, Rue Ch. Pauc, la chantrerie, 44306 Nantes Cedex 3, Nantes, France phone: +33.2.40.68.30.47, fax: +33. 2.40.68.32.32 e-mail: Patrick.Lecallet@univ-nantes.fr

Publikacja [8] to kolejna próba wykonania badań analogicznych jak w publikacjach [3] i [7]. Autorzy nieco więcej miejsca poświęcają wskaźnikowi subiektywnemu zebranemu dla 14 obserwatorów (z badań wyeliminowano 3 obserwatorów prezentujących inną niż inni ocenę obrazów – świadczy to tylko o mocno niepewnej metodyce wskaźników subiektywnych i związanej z nimi sporej dowolności). Testowane wskaźniki 2D (SSIM, UQI, C4, RRIQA) podobnie jak w innych badaniach nie nadawały się do oceny obrazów 3D. Próba wspomoczenia ich dołączonym wskaźnikiem oceniającym różnice w obrazie lewym i prawym również nie przyniosło poprawy sytuacji.

#### 4.9 Quality of Experience Assessment for Stereoscopic Images

Feng Qi<sup>1</sup>, Tingting Jiang<sup>2,3</sup>, Siwei Ma<sup>2,3</sup>, Debin Zhao<sup>1</sup> 1. School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China 2. National Engineering Lab for Video Technology 3. Key Lab.

of Machine Perception(MoE) School of EECS, Peking University, Beijing, China fqj@jdl.ac.cn, ttjiang@pku.edu.cn, swma@jdl.ac.cn, dbzhao@jdl.ac.cn

Podobny charakter jak publikacje [2] i [5]. Zdefiniowany współczynnik wymaga oczywiście dalszych testów.

#### *4.10 A comprehensive database and subjective evaluation methodology for quality of experience in stereoscopic video*

L. Goldmann, F. De Simone, and T. Ebrahimi, "EI, 3DIP and Applications, 2010

Publikacja opisuje przygotowaną bazę danych obrazów i video 3D, która może być używana do badania algorytmów (<http://mmspg.epfl.ch/3dvqa>).

#### *4.11 Algorithmic assessment of 3D quality of experience for images and videos*

Anish Mittal, Anush K. Moorthy, Joydeep Ghosh and Alan C. Bovik Dept. Of Electrical and Computer Engineering, The University of Texas at Austin, Austin, Texas - 78712.

Podobny charakter jak publikacje [2] i [5]. Zdefiniowany współczynnik wymaga uczenia co raczej ograniczy jego możliwości wykorzystania.

#### *4.12 Stereoscopic video quality assessment model based on spatial-temporal structural information*

Jingjing Han, Tingting Jiang, Siwei Ma National Engineering Lab. for Video Technology, Key Lab of Machine Perception (MOE), School of EECS, Peking University, Beijing, 100871 China {jjhan, ttjiang, swma}@jdl.ac.cn.

Podobny charakter jak publikacje [2] i [5]. Zdefiniowany współczynnik wymaga oczywiście dalszych testów.

#### *4.13 Quality Assessment of Stereoscopic Images*

Alexandre Benoit,<sup>1</sup> Patrick Le Callet (EURASIP Member),<sup>1</sup> Patrizio Campisi (EURASIP Member),<sup>2</sup> and Romain Cousseau<sup>1</sup> <sup>1</sup>Ecole Polytechnique de l'Université de Nantes, IRCCyN, rue Christian Pauc, 44306 Nantes Cedex 3, France <sup>2</sup>Dipartimento di Elettronica Applicata, Università degli Studi Roma Tre, Via della Vasca Navale 84, 00146 Roma, Italy campisi@uniroma3.it.

Wnioski analogiczne jak w publikacjach [3] i [8]

#### *4.14 Objective No-Reference Stereoscopic Image Quality Prediction, Based on 2D Image Features and Relative Disparity*

Z. M. Parvez Sazzad,<sup>1</sup> Roushain Akhter,<sup>2</sup> J. Baltes,<sup>2</sup> and Y. Horita<sup>1</sup> <sup>1</sup>Graduate School of Science and Engineering, University of Toyama, Toyama 930-8555, Japan <sup>2</sup>Department of Computer Science, University of Manitoba, Winnipeg, MB, Canada R3T 2N2.

Publikacja prezentuje podobną próbę rozwiązania problemu oceny obrazu 3D jak publikacje [3] i [8]. W tym wypadku jednak autorzy twierdzą, że osiągnęli sukces w postaci wysoko jakościowego współczynnika. Należy jednak zauważyć, że przetestowany zestaw obrazów ograniczał się wyłącznie do obrazów zakłóconych wyłącznie kompresją jpeg. Oznacza to, że wskaźnik raczej nie będzie dobrze funkcjonował w odniesieniu do innych zakłóceń i będzie mało przydatny.

#### *4.15 3D-DCTbased perceptual quality assessment of stereo video*

Lina Jin, Atanas Boev, Atanas Gotchev, Karen Egiazarian Department of Signal Processing, Tampere University of Technology (49\_Jin\_et\_al\_Sept2011.pdf)

Algorytm oceny bazuje na dyskretnej transformacji kosinusowej. Autorzy twierdzą że ich miara osiąga lepsze wyniki od wszystkich miar 2D, a także od miar 2D połączonych z oceną różnic obrazów lewego i prawego.

#### 4.16 Paired comparison-based subjective quality assessment of stereoscopic images

Jong-Seok Lee Lutz Goldmann Touradj Ebrahimi (37\_lee\_mta12\_online.pdf)

W publikacji zdefiniowano metodologię wyznaczania wskaźników subiektywnych. Ponieważ celem naszych badań są wskaźniki obiektywne tematyka publikacji ma mniejsze zastosowanie (wobec faktu posiadania już baz danych z zarejestrowaną metryką subiektywną dostępną do porównań).

Koniec

<b>Imię i nazwisko</b>	<b>Podpis</b>
Dr inż. Marek Kamiński	
Dr inż. Rafał Kotas	
Dr inż. Paweł Marciniak	
Dr inż. Bartosz Sakowicz	