

## BADANIA SUBIEKTYWNE JAKOŚCI QoE DLA USŁUG TELEWIZJI TRÓJWYMIAROWEJ

**Streszczenie:** Kolejnym etapem w rozwoju technik multimedialnych jest konstrukcja systemów umożliwiających użytkownikom końcowym percepcję efektu trójwymiarowego (3D). Prace naukowo-badawcze ukierunkowane na jakość postrzeganą przez użytkownika końcowego (ang. *Quality of Experience*, QoE) są w początkowej fazie - głównie ze względu na złożoność procesu ludzkiej percepcji scen trójwymiarowych. W tym artykule przedstawiono stan wiedzy w zakresie badań jakości dla usług telewizji 3D.

### 1. WSTĘP

Ostatni rok w dziedzinie technik multimedialnych został zdominowany przez rozwój i wdrożenia nowoczesnych usług oraz urządzeń umożliwiających użytkownikom końcowym percepcję głębi w wyświetlanych obrazach. Czołowi producenci wprowadzili na rynek wyświetlacze i telewizory 3D, smartfony i aparaty fotograficzne z podwójnymi obiektywami i wyświetlaczami autostereoskopowymi, a także pierwsze komputery przenośne z matrycami umożliwiającymi prezentację treści trójwymiarowej. Coraz łatwiej dostępne są również treści 3D – rozwijane są serwisy takie jak YouTube-3D, Video-on-Demand 3D, a także kanały telewizyjne [1].

Wydaje się, że obecny poziom rozwoju w tej dziedzinie i skala w jakiej jest finansowany zagwarantuje sukces. Nie udało się go osiągnąć w latach '50 ubiegłego wieku, kiedy po raz pierwszy prezentowano seanse 3D szerszej publiczności. Producenci wycofali się po kilkunastu miesiącach prosperity ze względu na spadek zainteresowania wywołany niskim komfortem widzów spowodowanym niezadowalającą jakością obrazu w trakcie projekcji [2].

Uwzględniając obecny stan zaawansowania technik multimedialnych *sensu largo*, nie ulega wątpliwości, że wysoka jakość obrazu jest punktem wyjścia przy podejmowaniu przez użytkowników końcowych decyzji o zakupie urządzeń. Dodatkowym wyzwaniem jest innowacyjność – dla wielu potencjalnych klientów wrażenie głębi wywoływane w sposób sztuczny jest doświadczeniem nowym, nieporównywalnym. Badania jakości obrazów 3D postrzeganej przez użytkowników (QoE) są przedmiotem ostatnich badań wielu zespołów naukowych na świecie – w czerwcu 2010 roku w ramach organizacji Video Quality Experts Group (VQEG) powstała nowa grupa „3DTV” [3, 4].

W tym artykule przedstawiono stan wiedzy w zakresie badań subiektywnych jakości QoE dla usług telewizji 3D. Zawiera on także próbę systematyzacji kryteriów oceny treści 3D przez użytkowników końcowych. Celem tej systematyzacji jest również identyfikacja pro-

blemów pojawiających się w trakcie przygotowywania testów subiektywnych dla telewizji 3DTV.

Struktura artykułu jest następująca: w sekcji 2 po krótko opisano istotę percepcji głębi rezygnując z analizy budowy układu wzrokowego (ang. *human visual system - HVS*). Architekturę systemu telewizji trójwymiarowej (3DTV) przybliżono w sekcji 3, a następnie w sekcji 4 wskazano na zniekształcenia w postaci artefaktów. W sekcji 5 omówiono zagadnienia związane z badaniami subiektywnymi QoE.

### 2. ISTOTA PERCEPCJI GŁĘBI

Anatomicznym uwarunkowaniem umożliwiającym człowiekowi percepcję głębi jest rozstaw gałek ocznych, który średnio jest równy 63 mm [5]. Najzasobniejszym źródłem informacji o głębi jest stereopsja, czyli proces dokonywany przez mózg polegający na specjalnym scaleniu obrazów odbieranych przez oczy. Naturalnie widzenie przestrzenne jest możliwe dzięki ściślejszej współpracy mechanizmów okoruchowych - akomodacji i konwergencji oraz stereopsji. Niemniej jednak większość z obecnie dostępnych stereoskopowych technik prezentacji nie respektuje powiązania tych mechanizmów zmuszając obserwatora do skupienia wzroku na z góry ustalonej odległości bez uwzględnienia faktycznego położenia obserwowanego obiektu w przestrzeni wynikającego z informacji stereopsyjnych [2].

Pozostałymi wskazówkami dla percepcji przestrzeni są: wskazówki „obrazowe” (okluzja, wysokość obiektów w polu widzenia, perspektywa geometryczna, relatywny rozmiar obiektów, cienie), paralaksa ruchowa (ang. *motion parallax*), a także tzw. funkcje kognitywne takie jak wnioskowanie, oczekiwanie, pamięć czy wiedza o obiektach [6, 7]. O ilości informacji zawartych w tych wskazówkach świadczy fakt, że zasłaniając jedno z oczu, wciąż jest możliwa percepcja przestrzeni.

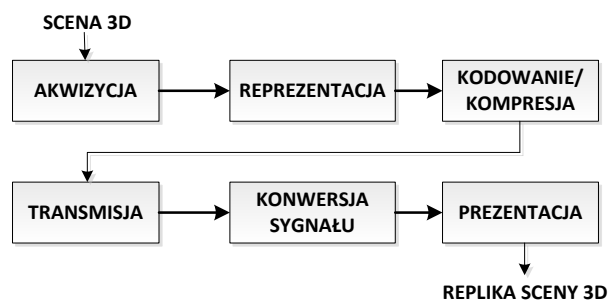
Ogromne znaczenie dla postrzegania głębi ma odległość obserwującego od obserwowanego obiektu. Należy zaznaczyć, że dla odległości większych od około 100 m, głębia jest rozpoznawana jedynie za pomocą wskazówek obrazowych i paralaksy ruchowej [8].

### 3. BUDOWA SYSTEMU TELEWIZJI 3DTV

System telewizji trójwymiarowej składa się sześciu bloków funkcjonalnych, które zostały przedstawione na rysunku 1. Na każdym etapie tego łańcucha sygnał ulega modyfikacjom, a to powoduje zniekształcenia (artefakty) wpływające na ostateczną ocenę jakości QoE przez użytkownika.

Informacje o rzeczywistej, trójwymiarowej scenie są rejestrowane w procesie akwizycji (pobierania) treści (ang. *content generation*). Jest on realizowany na trzy sposoby:

- za pomocą kamer z dwoma obiektywami (w tym przypadku znaczenie ma nie tylko rozstaw obiektywów, ale także ich położenie względem siebie – konfiguracja zbieżna lub równoległa [9]) bądź za pomocą wielu kamer,
- konwencjonalnej kamery wyposażonej w kamerę rejestrującą głębię [10, 11],
- za pomocą algorytmów konwersji obrazów dwuwymiarowych na obrazy trójwymiarowe [12].



Rys. 1. System 3DTV

Sposób akwizycji determinuje reprezentację danych, a co za tym idzie także sposób kodowania i kompresji. Właściwie można wyróżnić trzy główne grupy relacji:

1) wideo wielowidokowe (ang. *multiview video*) – gdy dysponujemy więcej niż dwoma strumieniami wideo zawierającymi treść pobraną z różnych punktów obserwacyjnych. Do kompresji wykorzystywane są między innymi algorytmy H.264/MVC [13].

2) *video+depth* – gdy oprócz sekwencji 2D dysponujemy mapą głębi korespondującą z zawartością tej sekwencji. W tym przypadku stosowane są standardy MPEG ISO/IEC 23002-3 (MPEG-C).

3) dynamiczne siatki 3D – gdy geometria sceny jest reprezentowana przez modele siatek 3D. Ten rodzaj reprezentacji jest efektywny dla nagrań syntetycznych. Metody kompresji zostały usystematyzowane w rozszerzeniu do standardu MPEG-4 [14].

Transmisja strumieni może być zrealizowana za pomocą protokołów RTP/UDP, RTP/TCP, RTP/DCCP (dwa rodzaje kontroli natłoku (ang. *congestion control*): TCP-like congestion control (IETF RFC 4341) oraz TCP-friendly rate control (TFRC) (IETF RFC 4342)) [15].

Ostatnim, najbliższym użytkownikowi końcowemu, blokiem funkcjonalnym systemu 3DTV jest prezentacja treści 3D. Z wyjątkiem napotykaną na ograniczenia technologiczne holografii i metod wolumetrycznych, wszystkie dotychczas zaproponowane metody prezentacji obrazu trójwymiarowego bazują na separacji kanałów – sekwencji przeznaczonych odpowiednio dla lewego i prawego oka, nieznacznie różniących się od siebie. Ze względu na ogromną popularność aplikacji na rynku wyróżnić należy dwie metody kierunkowe: wymagające użycia specjalnych okularów (metody stereoskopowe)

oraz niewymagające takiego wyposażenia (metody autostereoskopowe).

W przypadku metod stereoskopowych separacja kanałów może być zrealizowana za pomocą polaryzacji lub multipleksacji czasowej (zastosowanie anaglifów znacząco obniża jakość odzwierciedlenia kolorów [16]) [17]. Z kolei w przypadku autostereoskopowych wykorzystywane są bariery paralaksy (ang. *parallax barrier*) bądź powłoki mikrosoczewkowe (ang. *lenticular sheet*) [18].

#### 4. ARTEFAKTY

Artefakty widoczne dla użytkownika końcowego mogą zostać sklasyfikowane według dwóch kategorii: 1) źródła, z którego pochodzą (odpowiednie bloki funkcjonalne systemu), 2) ze względu na obszar oddziaływania – wpływają na strukturę obrazu, kolor, odzwierciedlenie dynamiki obrazu oraz elementy odpowiedzialne za wywołanie wrażenia głębi – unikatowe dla obrazów trójwymiarowych [19]. W ramach drugiej kategorii w literaturze wskazuje się na następujące czynniki powodujące dyskomfort: zakrzywienie płaszczyzny w głębi (ang. *depth plane curvature*) [20], zniekształcenia geometryczne (trapezowe) obrazu, zbieg perspektyw (ang. *keystone distortion*), nielinowość głębi (ang. *depth non-linearity*), efekt skracania głębi (ang. *shear distortion*), efekt teatru kukiełek (ang. *puppet theatre effect*), efekt tekturowy (ang. *cardboard effect*) [21], przesłuchy (ang. *crosstalk*) [22, 23, 31], efekt płotu sztachetowego (ang. *picket fence effect*), drżenie obrazu (ang. *image flipping*) [24], a także propagację błędów spowodowaną przez utratę pakietów [2, 25].

Efektem, który ma największy wpływ na obniżenie jakości obrazu i komfortu oglądania są przesłuchy, czyli zniekształcenia spowodowane niedoskonałością separacji kanałów dla lewego i prawego oka [29]. Organizacja VQEG w ramach sekcji 3DTV rozpoczęła swoje prace od przygotowania precyzyjnego scenariusza do przeprowadzania badań QoE (tzw. test-plan) ukierunkowanych na analizę wpływu przesłuchów na percepcję głębi i oceny użytkowników.

#### 5. BADANIA SUBIEKTYWNE – QoE-3D

Przeprowadzanie testów subiektywnych jest szczególnie uzasadnione w przypadku usług 3DTV, ponieważ dostarczają one informacji o tym jak użytkownicy postrzegają i oceniają jakość prezentowanych sekwencji trójwymiarowych nie posiadając tak bogatego doświadczenia (porównania) jak przy ocenie usług 2D [26].

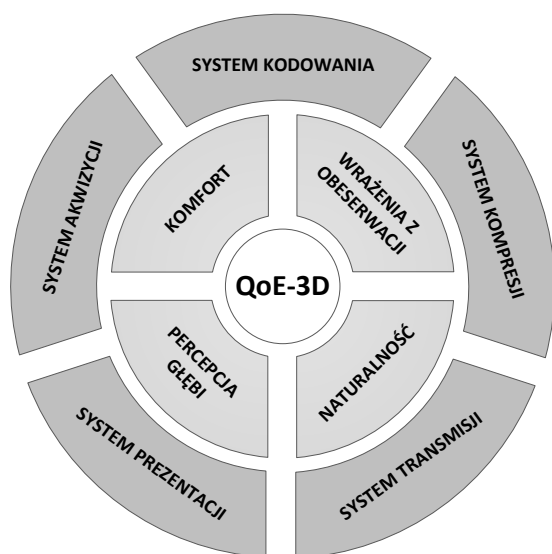
##### 5.1. Metoda badań QoE – nowe wymagania

Metoda przeprowadzania testów subiektywnych dla obrazów trójwymiarowych została uregulowana w rekomendacji ITU-R BT.1438, która istotnie nawiązuje do rekomendacji ITU-R BT.500. Niestety wiele szczegółów specyficznych dla badania jakości 3D nie znalazło miejsca w tym opracowaniu. W wyniku postępu w prowadzeniu badań tego typu, w [27] przedstawiono w sposób

uporządkowany propozycje zmian uwarunkowań technicznych.

Trudność w przygotowaniu precyzyjnych standardów wynika:

- ze złożoności systemu 3DTV (sekcja 3),
- z różnorodności urządzeń służących do prezentacji [17],
- z wieloaspektowości, wielowymiarowości jakości obrazów 3D jako konstrukcji psychologicznych [28, 29, 30],
- z integralności, wzajemnego wpływu poszczególnych czynników percepcyjnych (np. wraz ze wzrostem różnicy między obrazami stereopsyjnymi zwiększa się wrażenie głębi, ale zwiększają się także przesłuchy),
- z dodatkowych parametrów percepcyjnych (podsekcja 5.2).



Rys. 2. Koncepcja QoE-3D jako system-of-systems.

Na rysunku 2 schematycznie przedstawiono ideę wielowymiarowości badań jakości QoE dla multimediów 3D. Idea ta nawiązuje do koncepcji *system-of-systems*. Warstwę zewnętrzną stanowią czynniki wpływające na jakość QoE, wynikające wprost z elementów systemu 3DTV. Z kolei warstwę wewnętrzną stanowią kryteria oceny treści 3D przez użytkowników końcowych.

## 5.2. Subiektywne wizualne kryteria oceny treści 3D

Do przeprowadzenia testów subiektywnych oprócz przygotowania odpowiednio zniekształconych sekwencji należy skonstruować zestaw pytań skierowanych do testerów. Ze względu na ich opiniodawczy charakter, potencjalne odpowiedzi powinny zostać sformułowane zgodnie z wymaganiami skali Likerta, aby wyniki uzyskane od kolejnych obserwatorów były porównywalne. Jest to skala pięciostopniowa i pozwala uzyskać odpowiedź dotyczącą stopnia akceptacji danego zjawiska, ponieważ badany wybiera opcję, która najbardziej odpowiada jego odczuciom. Z pewnością konieczne jest przeprowadzenie badań pilotażowych, aby stwierdzić czy zadawane pytania mają charakter dyskryminacyjny i dostarczają wartościowych informacji.

Można wyróżnić kilka grup wyrażen nawiązujących do konkretnych wrażeń dających się określić jako kryteria oceny jakości. Brak precyzji jest konsekwencją bałaganu pojęciowego występującego w literaturze, którego niestety nie udało się uniknąć. Poniekąd jest on w pełni usprawiedliwiony specyfiką badań subiektywnych związanych z 3D – technologią wciąż innowacyjną i prężnie rozwijającą się.

Pierwsza grupa dotyczy subiektywnego wrażenia bycia w danym miejscu czy środowisku, nawet fizycznie tkwiąc w zupełnie innym. Gdy to wrażenie jest intensywne, obserwatorzy bardziej koncentrują się na kreacji dostarczonej przez medium niż na swoim fizycznym otoczeniu. W literaturze anglojęzycznej do opisu używane są następujące określenia: *presence*, *immersion*, *involvement*, *sense of „being there”*.

Do drugiej grupy należą przymiotniki *naturalness*, *realism* oraz ich synonimy. Należy je rozumieć jako autentyczność reprezentacji rzeczywistości, obserwatorzy oceniają proporcje ludzi i obiektów w przestrzeni (głębi), unikając oceny zawartości merytorycznej sekwencji.

Komfort obserwacji jest najczęściej określany jako: *viewing experience*, *visual comfort*, *“eye-strain”*, *„pleasing” image*. Do tej grupy należą zatem także określenia związane ze zmęczeniem wzroku, uczuciem nudności, zawrotami głowy.

Ostatnią grupę można zawrzeć w określeniu „ogólna ocena jakości”. Obserwatorzy oceniają globalnie jakość obrazu z uwzględnieniem efektu głębi.

W [30] oprócz czynników technologicznych (format danych, kompresja, reprezentacja głębi, przesłuchy, zniekształcenia) pośrednio wpływających na jakość QoE wymieniane są czynniki bezpośrednio oddziałujące na percepcję takie jak wpływ głębi, stałość jakości, pole widzenia, wrażenie z obserwacji, obecność (ang. *presence*).

## 6. PODSUMOWANIE

W artykule przedstawiono obecny stan wiedzy w zakresie badań jakości postrzeganej przez użytkowników QoE dla usług telewizji trójwymiarowej. Wskazano na wielowymiarowość relacji pomiędzy poszczególnymi elementami wpływającymi na percepcję, jak również na szereg problemów wymagających szczegółowych badań. Wyznaczono kierunek najbliższych prac.

## 7. PODZIĘKOWANIA

Artykuł powstał w ramach Programu Operacyjnego Innowacyjna Gospodarka - projekt nr POIG.01.01.01-00-045/09-00 „Inżynieria Internetu Przyszłości” sfinansowanego przez Ministerstwo Nauki i Szkolnictwa Wyższego z Europejskiego Funduszu Rozwoju Regionalnego.

## SPIS LITERATURY

- [1] B. Bing, *3D and HD Broadband Video Networking*, Artech House Publishers, 2010

- [2] O. Schreer, P. Kauff, T. Sikora, *3D Videocommunication: Algorithms, concepts and real-time systems in human centred communication*, John Wiley & Sons, 2005
- [3] <http://www.its.bldrdoc.gov/vqeg/>
- [4] A. van Moorsel, "Metrics for the Internet Age: Quality of Experience and Quality of Business", *Fifth Performability Workshop*, September 16, 2001, Erlangen, Germany
- [5] N. A. Dodgson, "Variation and extrema of human interpupillary distance", *Stereoscopic Displays and Virtual Reality Systems XI*, 5291(1), pp. 36–46, 2004
- [6] P.J.H. Seuntjens, *Visual Experience of 3DTV*, PhD dissertation
- [7] J. Mansson, *Stereovision, A model of human stereopsis*, Technical report, 1998.
- [8] B.A. Wandell, *Foundations of vision*, Sinauer Associates, Inc. Sunderland, Massachusetts, USA, 1995
- [9] A. J. Woods, T. Docherty, R. Koch, "Image distortions in stereoscopic video systems", *Stereoscopic Displays and Applications IV*, 1915(1), pp. 36–48, 1993.
- [10] M. Kawakita, K. Iizuka, T. Aida, H. Kikuchi, H. Fujikake, J. Yonai, K. Takizawa, "Axi-Vision Camera (Real-Time Distance-Mapping Camera)", *Applied Optics*, Vol. 39, Issue 22, pp. 3931–3939 (2000)
- [11] <http://web.archive.org/web/20050211113338/3dvsystems.com/news/release1.html>
- [12] C. Fehn, P. Kauff, M. Op De Beeck, F. Ernst, W. IJsselstein, M. Pollefeys, L. Van Gool, E. Ofek, I. Sexton, "An Evolutionary and Optimised Approach on 3D-TV", *Proceedings of International Broadcast Conference*, 2002, pp. 357–365
- [13] I.E. Richardson, *The H.264 Advanced Video Compression Standard*, Wiley, 2010
- [14] MPEG-4 Animation Framework eXtension (AFX), ISO/IEC 14496-16 2<sup>nd</sup> ed. [online] – dostępny: [www.mpeg-3dgc.org](http://www.mpeg-3dgc.org)
- [15] G.B. Akar, A.M. Tekalp, C. Fehn, M.R. Civanlar, "Transport methods in 3DTV – a survey", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17 issue 11, pp. 1622–1630, Nov. 2007
- [16] E. Dubois, "A projection method to generate anaglyph stereo images", *ICASSP '01: Proceedings of the Acoustics, Speech, and Signal Processing, on IEEE International Conference*, pp. 1661–1664, Washington, DC, USA, 2001, IEEE Computer Society.
- [17] M. Konrad, J. Halle, "3-D displays and signal processing", *Signal Processing Magazine, IEEE*, 24(6), pp. 97–111, nov. 2007
- [18] N.A. Dodgson, "Autostereoscopic 3D displays", *Computer*, 38(8), pp. 31–36, aug. 2005.
- [19] A. Boev, D. Hollosi, A. Gotchev, K. Egiazarian, "Classification and simulation of stereoscopic artifacts in mobile 3DTV content", *Proceedings of the SPIE 7237*, 72371F, 2009
- [20] A.J. Woods, T. Docherty, R. Koch, "Image distortions in stereoscopic video systems", *Stereoscopic Displays and Applications IV*, 1915(1), pp. 36–48, 1993
- [21] H. Yamanoue, "The relation between size distortion and shooting conditions for stereoscopic images", *Journal of the SMPTE*, 106(4), pp. 225–232, 1997
- [22] S. Pastoor, Human factors of 3D imaging: Results of recent research at Heinrich-Hertz-Institut, *ASIA Display '95 Conference*, pp. 66–72, Berlin, 1995, *IEEE Computer Society*.
- [23] A. J. Woods, "Understanding Crosstalk in Stereoscopic Displays", *3DSA (Three-Dimensional Systems and Applications) conference*, Tokyo, Japan, 19–21 May 2010
- [25] A. Boev, D. Hollosi, and A. Gotchev, "Software for simulation of artifacts and database of impaired videos", *Mobile3DTV Project report*, dostępny: <http://mobile3dtv.eu>
- [26] J. Freeman, S. Avons, Focus group exploration of presence through advanced broadcast services, *Proceedings of the SPIE*, 2000
- [27] W. Chen, J. Fournier, M. Barkowsky, Patric Le Callet, "New requirements of subjective video quality assessment methodologies for 3DTV", *Fifth International Workshop on Video Processing and Quality Metrics for Consumer Electronics - VPQM 2010*.
- [28] L. Meesters, Predicted and Perceived Quality of Bit-Reduced Gray-Scale Still Images, *Eindhoven University Press Facilities*, 2002
- [29] L.M. Meesters, W.A. IJsselstein, P.J.H. Seuntjens, "A survey of perceptual evaluations and requirements of three-dimensional TV", *Circuits and Systems for Video Technology, IEEE Transactions on*, vol.14, no.3, pp. 381–391, March 2004
- [30] A. Benoit, P. Le Callet, P. Campisi, R. Cousseau, "Quality Assessment of Stereoscopic Images", *EURASIP Journal on Image and Video Processing*, vol. 2008
- [24] I. Sexton, P. Surman, Stereoscopic and autostereoscopic display systems, *Signal Processing Magazine, IEEE*, 16(3), pp. 85–99, May 1999
- [31] L. Xing, J. You, T. Ebrahimi, A. Perkis, "An objective metric for assessing quality of experience on stereoscopic images," *Multimedia Signal Processing (MMSP), 2010 IEEE International Workshop on*, vol., no., pp. 373–378, 4–6 Oct. 2010