



## 2B Video ve Derinlik Haritası Temelli 3B Videolarda Yapısal Karmaşıklık Tabanlı Video Kalite Değerlendirmesi

### Color Plus Depth Map Based 3D Video Quality Assessment Based on Structural Complexity

Mustafa Yasin ERTEN, Gökçe NUR YILMAZ

*Kırıkkale Üniversitesi, Elektrik ve Elektronik Mühendisliği Bölümü, 71450 Kırıkkale, TÜRKİYE*

Başvuru/Received: 20/11/2017

Kabul/Accepted: 05/01/2017

Son Versiyon/Final Version: 29/01/2018

#### Öz

Günümüz teknolojilerindeki gelişmeler, yüksek bant genişliğine sahip haberleşme kanallarının varlığı ve büyük video servis sağlayıcılarının tüketici piyasasına girmesi 3 Boyutlu (3B) video teknolojilerinin hızlı gelişim göstermesine neden olmuştur. 3B video teknolojilerine geri bildirim oluşturabilmesi ve bu gelişime destek olması dolayısı ile 3B Video Kalite Değerlendirmesi (VKD) günümüzde 3B video servis sağlayıcıları için oldukça bir ihtiyaç haline gelmiştir. Literatürde zaman alıcı ve maliyetli öznel testlerin yerini alabilecek, verimli, güvenilir ve yaygın olarak kullanılan nesnel bir 3B VKD metriği bulunmadığından 3B VKD yerine kullanılan 2 Boyutlu (2B) VKD metriklerinin çoğu Tam Referanslı (TR) ya da Azaltılmış Referanslı (AR) metriklerdir. Bu metriklerin hem gerçek zamanlı kalite ölçümü gerçekleştirmelerindeki eksiklikten hem de 3B videoların iletimi için yüksek bant genişliğine ihtiyaç duyulmasından ötürü bu yöntemlerin verimli olmadığı anlaşılmıştır. Dolayısıyla orijinal videodan ek bir bilgi gerekmeden alıcı tarafta sadece sıkıştırılmış 3B videoyu kullanarak değerlendirme yapan referanssız bir 3B VKD metriğinin kullanılması elzem olmuştur. Bu çalışmada, IGS ile oldukça bağlantılı olan, 3B videolardaki objelerin kenar bilgileri kullanılarak oluşturulan Yapısal Karmaşıklık (YK) temelli bir Referanssız 3B VKD metriği geliştirilmiştir. Geliştirilen bu metrik kullanılarak elde edilen verimli sonuçlar metriğin 3B video servislerinin geliştirilmesine büyük katkı sağlayabileceğini göstermiştir.

#### Anahtar Kelimeler

“3B video, video kalite değerlendirmesi, yapısal karmaşıklık”

#### Abstract

3-Dimensional (3D) video technologies present rapid improvement due the development of today's technologies, the presence of the communication channels having high bandwidth capacities, and the entrance of the powerful video service providers into the consumer electronics market. 3D Video Quality Assessment (VQA) becomes a must for the 3D video service providers via feedback it provides to the 3D video technologies and its support to this improvement. There is currently lack of an efficient, reliable, and commonly utilized objective 3D VQA metric that is preferable than the time-consuming and costly subjective tests. The 2-Dimensional (2D) VQA metrics used instead of the 3D VQA metrics to fulfill this lack are mostly Full Reference (FR) metrics or Reduced Reference (RR) metrics. These metrics are considered inefficient due to their insufficiency at online quality evaluation and requirement of high bandwidth capacity during the 3D video transmission. Therefore, there becomes a need for a No Reference (NR) 3D VQA metric that does not require any information related to the original video but only uses the compressed 3D video for the assessment. In this study, a NR 3D VQA metric which is based on Structural Complexity (SC) formed using edge information of the objects existing in a 3D video is developed. The efficient results obtained using the proposed metric prove that tremendous support can be provided to improve the 3D video services using the proposed metric.

#### Key Words

“3B video, video quality assessment, structural complexity”

## 1. GİRİŞ

Tasarlanan modern 3 Boyutlu (3B) sistem ve teknolojilerindeki ana amaç izleyicilere tatmin edici görsel deneyim kazandırmaktır. Bu deneyimi 3B Video Kalite Değerlendirme (VKD) yöntemleri ile güvenilir bir şekilde ölçmek, bu sistemlerin ve teknolojilerin gelişimi açısından oldukça önemlidir. Literatürde 3B VKD için, gerçek gözlemcileri kullanan öznel testler ve bilgisayar programları destekli matematiksel olarak ölçüm yapan nesnel metrikler bulunmaktadır. Öznel testler, bu testler sırasında gözlemcilerin görsel algılarını kalite değerlendirmesine etkili şekilde aktarabilmesi ve literatürde 3B VKD için yeterince güvenilir nesnel bir metriğin bulunmaması nedeniyle sıklıkla tercih edilmektedir. Fakat öznel testlerin test platformunun kurulmasındaki, izleyici seçimindeki ve uzun değerlendirme sürelerindeki zorluklar sebebiyle teknolojinin gelişmesine ve ticari bir ürün gelişimine negatif etkileri vardır. Bu nedenle güvenilir nesnel 3B VKD metriklerin geliştirilmesi oldukça önem arz etmektedir (Huynh-Thu vd., 2012) (Yılmaz vd., 2017) (Yılmaz vd., 2015).

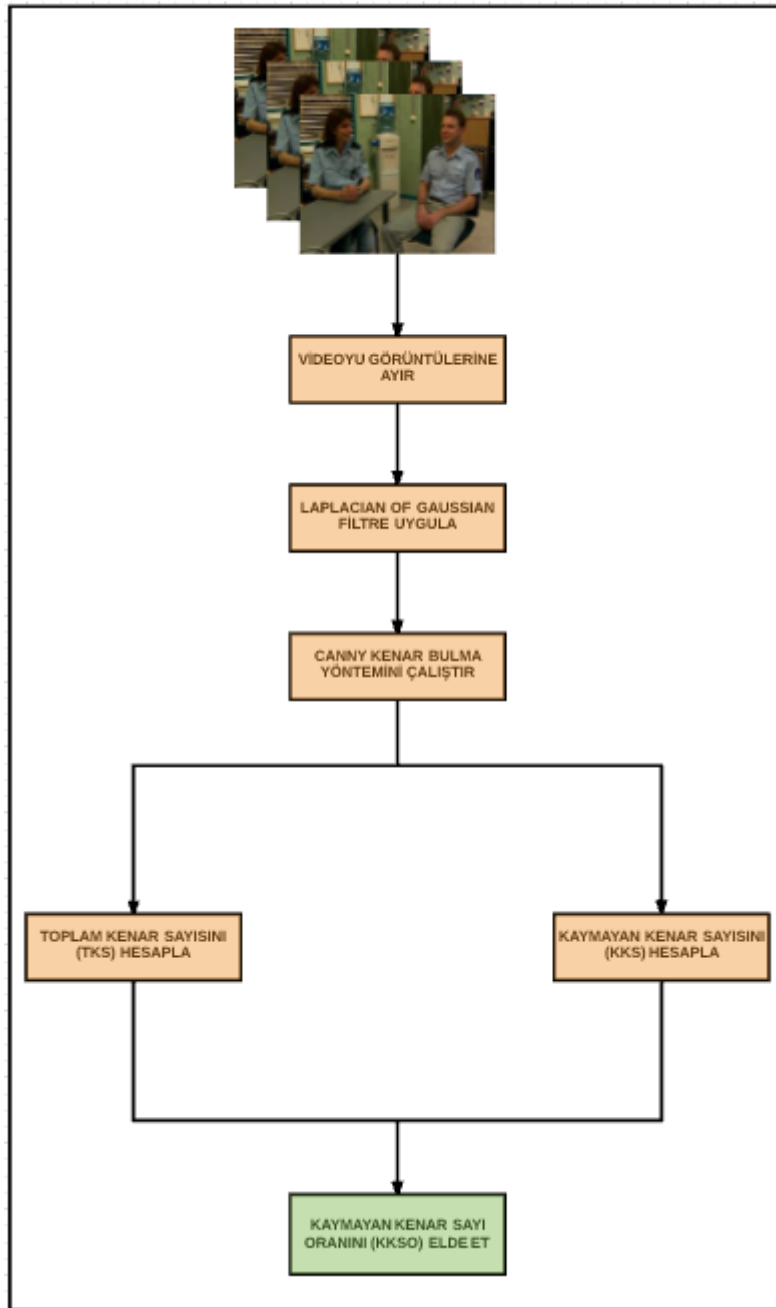
Nesnel 3B VKD metrikleri Tam Referanslı (TR), Azaltılmış Referanslı (AR)'lı ve referanssız metrikler olmak üzere üçe ayrılmaktadır. TR metrikler 3B VKD için kullanıcı tarafında orijinal videoya ihtiyaç duymaktadırlar. AR metrikler ise orijinal videodan çıkarılan bilgi ile 3B VKD gerçekleştirmektedirler. Bunların yanı sıra, referanssız metrikler ise orijinal video ile ilgili hiçbir bilgiye ihtiyaç duymadan 3B VKD'yi gerçekleştirmektedirler. Dolayısı ile referanssız metrikler özellikle online video iletiminde gayet verimlidirler (Yılmaz vd., 2017) (Yılmaz vd., 2015). Bu bilgiler ışığında bu çalışmada 3B VKD için referanssız metrik türü temel alınmıştır.

Video görsel deneyimi; kenar, arka plan ve özyapı gibi videonun yerel özelliklerine fazlasıyla bağlıdır. Bu özelliklerden kenar bilgisi genellikle resmin birçok özelliğini taşımaktadır ve literatürde gerçekleştirilen çalışmalar video kenar bilgilerinin videonun Yapısal Karmaşıklık (YK) ile yüksek uyumluluk gösterdiğini belirtmişlerdir (Fu vd., 2008) (Martini vd., 2012). Bu bilgiler ışığında bu çalışmada, 3B video görüntülerindeki objelerin kenar bilgileri ile bağlantılı YK özelliğine dayanan ve dolayısı ile İGS ile yüksek uyuma sahip bir metrik geliştirilmiştir. Bu metrik Yapısal Karmaşıklık Algoritması (YKA) olarak adlandırılmıştır. 2 Boyutlu (2B) video + Derinlik Haritası (DH) temelli 3B videonun diğer gösterim şekillerine göre avantajları (De Silva vd., 2012) dolayısı ile YKA geliştirilirken bu gösterim şekli kullanılmıştır.

Bu makale dört bölümden oluşmaktadır. İkinci bölümde kenar bilgisi çıkarımı tabanlı YKA tanıtılmıştır. Üçüncü bölümde algoritma sonucu elde edilen sonuçlar değerlendirilmiştir. Dördüncü bölümde ise gelecekte yapılabilecek çalışmalar irdelenmiştir.

## 2. YAPISAL KARMAŞIKLIK ALGORİTMASI

Tasarlanan YKA'nın blok diyagramı Şekil 1'de gösterilmiştir. Şekilden de görüleceği üzere YKA algoritma tasarımının ilk adımında 2B videolar görüntülere ayrılmıştır. Gürültünün video kalitesinin ve dolayısı ile videonun YK' sının ölçümünün düşmesinde en önemli faktör olduğu öngörüsü ile önerilen YKA'yı tasarlarken ki ikinci adımımız 2B videoları gürültüden ayırmaktır. Şekil 1'den de gözlemlenebileceği gibi, tasarlanan algoritmada, gürültüden ayırım için, 2B video görüntülerinin düşük frekanslı bileşenlerini filtreleyen ve yüksek frekanslı ve yapısal bilgileri barındıran bileşenlerini geçiren Laplacian of Gaussian (LoG) filtre kullanılmıştır (De Silva vd., 2012). Bu filtrenin tercih edilme sebebi daha sonra hakkında detaylı bilgi vereceğimiz kenar bulma algoritması (Canny algoritması) ile uyumlu çalışmasıdır (Fu vd, 2008).



Şekil 1: YKA'nın blok diyagramı

LoG filtre görüntünün ikinci dereceden uzamsal türevini hesaplar. Bu türev görüntüde sabit piksel yoğunluğunun olduğu alanlarda yani yoğunluk gradyanının sıfır olduğu yerlerde LoG sonucu sıfır olduğu anlamına gelir. Yoğunlukta değişikliğin olduğu durumlarda ise LoG sonucu görüntünün karanlık bölümlerinde pozitif, aydınlık olduğu bölümlerde ise negatiftir. Denklem(1)'de LoG filtrenin formülasyonu gösterilmiştir (Huynh-Thu vd., 2012):

$$LoG(i, j) = \frac{1}{\pi\sigma^4} \left[ 1 - \frac{i^2 + j^2}{2\sigma^2} \right] e^{-\frac{i^2 + j^2}{2\sigma^2}} \quad (1)$$

Burada,  $LoG(i, j)$  i ve j piksel koordinatlarındaki LoG filtre cevabını,  $\sigma$  ise Gauss dağılımının standart sapmasını göstermektedir. Denklem (1)'de  $LoG(i, j)$  daha sonra her bir görüntüye aşağıdaki denklemde verilen konvolüsyon işlemi yolu ile uygulanmıştır:

$$\beta_i = LoG(i, j) * R(i, j) \quad (2)$$

Burada,  $\beta_i$  LoG filtresi uygulanan görüntüleri,  $R(i, j)$  ise video boyunca bulunan her bir görüntüyü ifade etmektedir.

YKA tasarımının üçüncü aşaması olarak LoG operatörüyle filtrelenen görüntülerin Canny kenar bulma algoritması yardımıyla kenar bilgileri çıkarılmıştır. Video görüntülerindeki nesnelerin sınırlarını karakterize eden kenar bilgisi, renkli özyapı dizisinin yapısal özelliğini temsil etmektedir. Canny kenar bulma algoritması, video görüntülerindeki kenar bilgisini temsil eden pikselleri bastırmadan bulan ve bu bulunan piksellere 1 değeri atayan diğerlerine ise 0 değeri atayan bir algoritmadır (JVSM). YKA tasarımı için, her bir  $\partial_i$  görüntüsünde atanan piksel değerleri toplanmıştır ve tüm videoya genişletilmiştir. Bu genişletilmiş toplam değeri Toplam Kenar Sayısı (TKS) olarak ifade edilmiştir. Denklem (3)'te TKS hesabının matematiksel ifadesi belirtilmektedir:

$$TKS = \sum_{i=1}^{GS} \sum_{k=1}^{M \times N} \partial_{i,k} \quad (3)$$

Burada,  $\partial_{i,k}$  her bir görüntüdeki 1 değeri atanan pikselleri belirtmektedir.  $M \times N$  görüntünün genişlikxyüksekliğini ifade etmektedir. GS görüntü sayısını ifade etmektedir dolayısı ile  $\partial_i$ 'lerin sayısına işaret etmektedir. YKA tasarımının dördüncü aşamasında, görüntünün orijinal kenar pozisyonlarını koruyan kenar bilgilerinin sayısının VKD için kritik öneme sahip olduğu öngörüsü ile birbirini takip eden görüntülerin Kaymayan Kenar Sayısı (KKS) şu şekilde bulunmuştur:

$$KKS = \sum_{i=2}^{GS} \sum_{k=1}^{M \times N} (\partial_{i,k} \cap \partial_{i-1,k}) \quad (4)$$

YKA tasarımının son aşamasında ise, KKS değerleri TKS değerleri ile normalize edilerek Kaymayan Kenar Sayısı Oranı (KKSO) elde edilmiştir:

$$KKSO = \frac{KKS}{TKS} \quad (5)$$

KKSO, YKA algoritması sonucu elde edilen değerdir.

### 3. SONUÇLAR VE TARTIŞMALAR

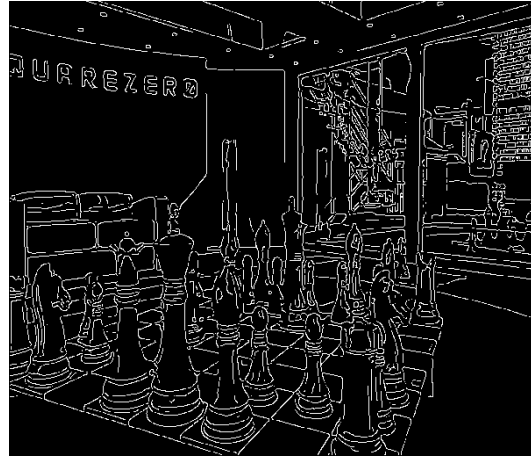
Bölüm 2'de bahsedilen YKA algoritmasının ilk iki adımı (LoG filtre ve Canny algoritması) kullanılarak, 2B video+DH bazlı orijinal (kodlanmamış) Chess, Interview, ve Windmill videolarının 2B videolarının ilk ve son görüntülerinin kenarlarının çıkarılmış hali Şekil 2 de resmedilmiştir. Şekilden de görüleceği üzere, Chess videosunun diğer videolara göre daha belirgin ve çok kenara sahiptir.

Bu çalışmadaki her bir videonun (Chess, Interview, ve Windmill) kodlanmalarında kullanılan bit oranlarına (768, 1024 ve 1536 kbps) göre ayrı ayrı hesaplanan KKSO değerleri Şekil 3'te tüm görüntüler için ayrı ayrı ve Şekil 4'te hepsi için birlikte gösterilmektedir. Bu videoları bu bit oranlarında kodlarken JSVM 9.13.1 kodek kullanılmıştır.

Şekil 3 ve 4 incelendiğinde, Chess videosu KKSO değeri en yüksek videodur. Fakat bit rate değişimi bu video için diğerlerine kıyasla daha az KKSO değişimi göstermektedir. Yine şekilden gözlemleneceği üzere, Interview videosunda ise bit rate oranı arttıkça KKSO değerlerinde yaklaşık aynı oranda artış yaşanmaktadır. Windmill videosu için KKSO değerlerinde 768 kbps ve 1024 kbps değerleri arasında ufak bir yükseliş yaşanırken 1024 kbps ve 1536 kbps değerleri arasında bir önceki değerine yaklaşık beş katı bir yükseliş gözlemlenmiştir.



(a)



(b)



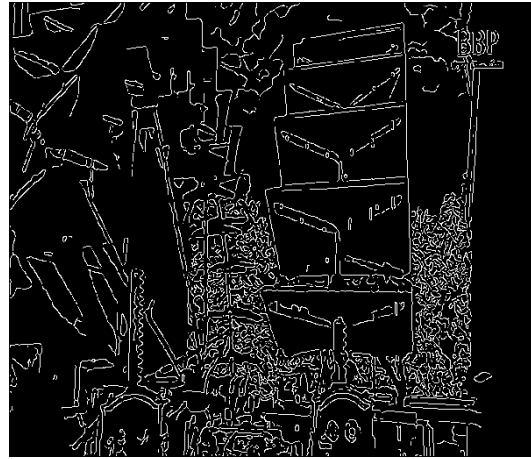
(c)



(d)

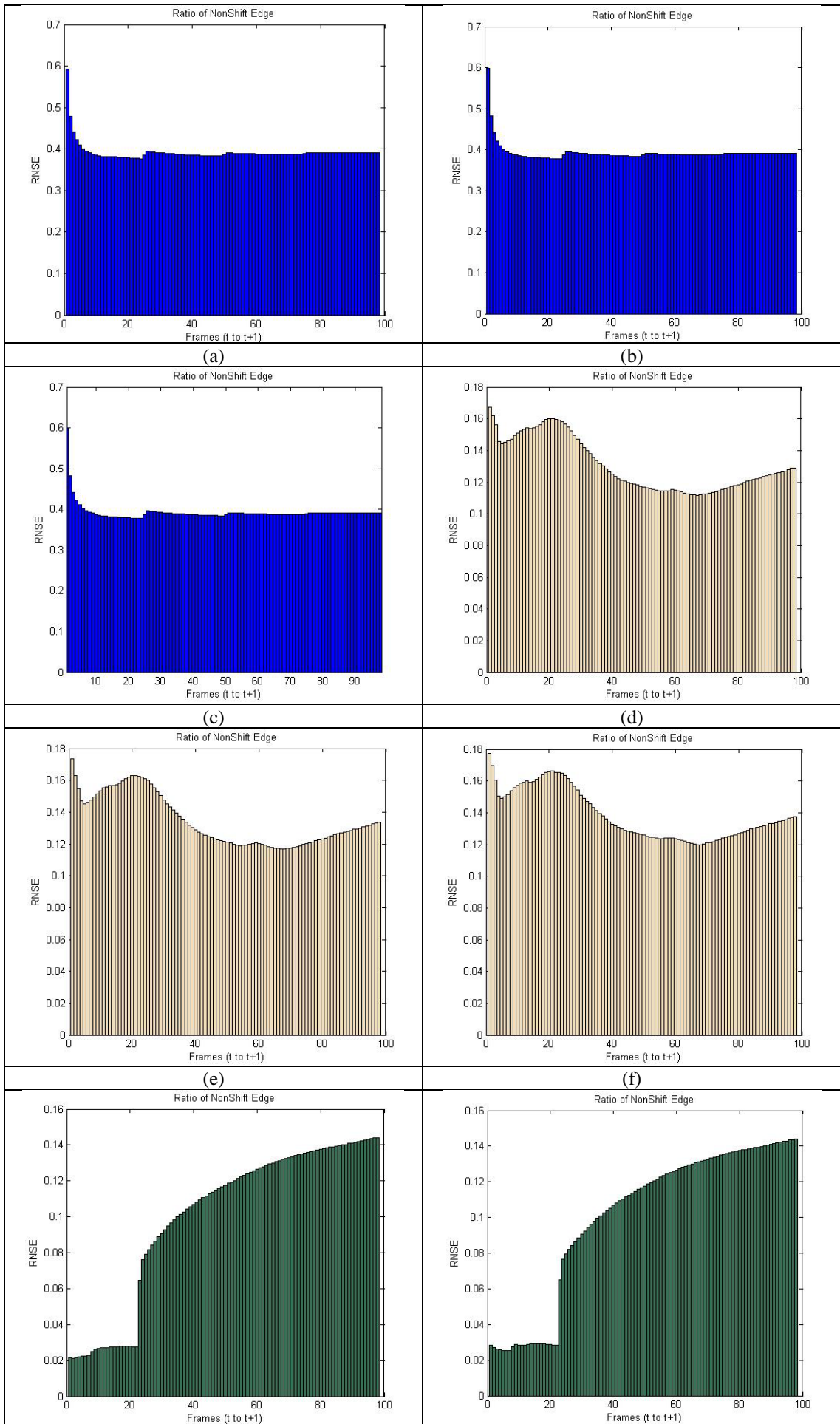


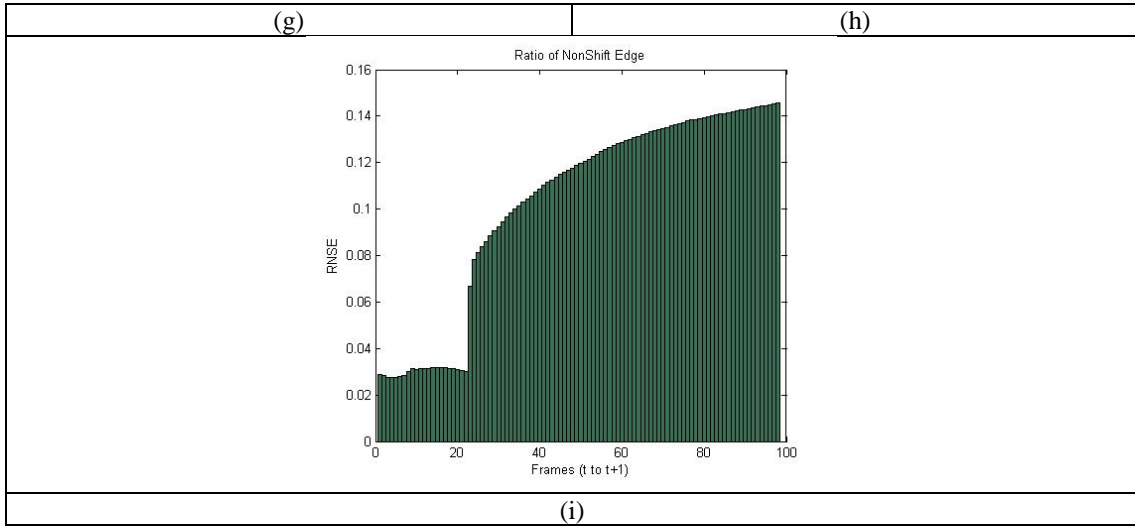
(e)



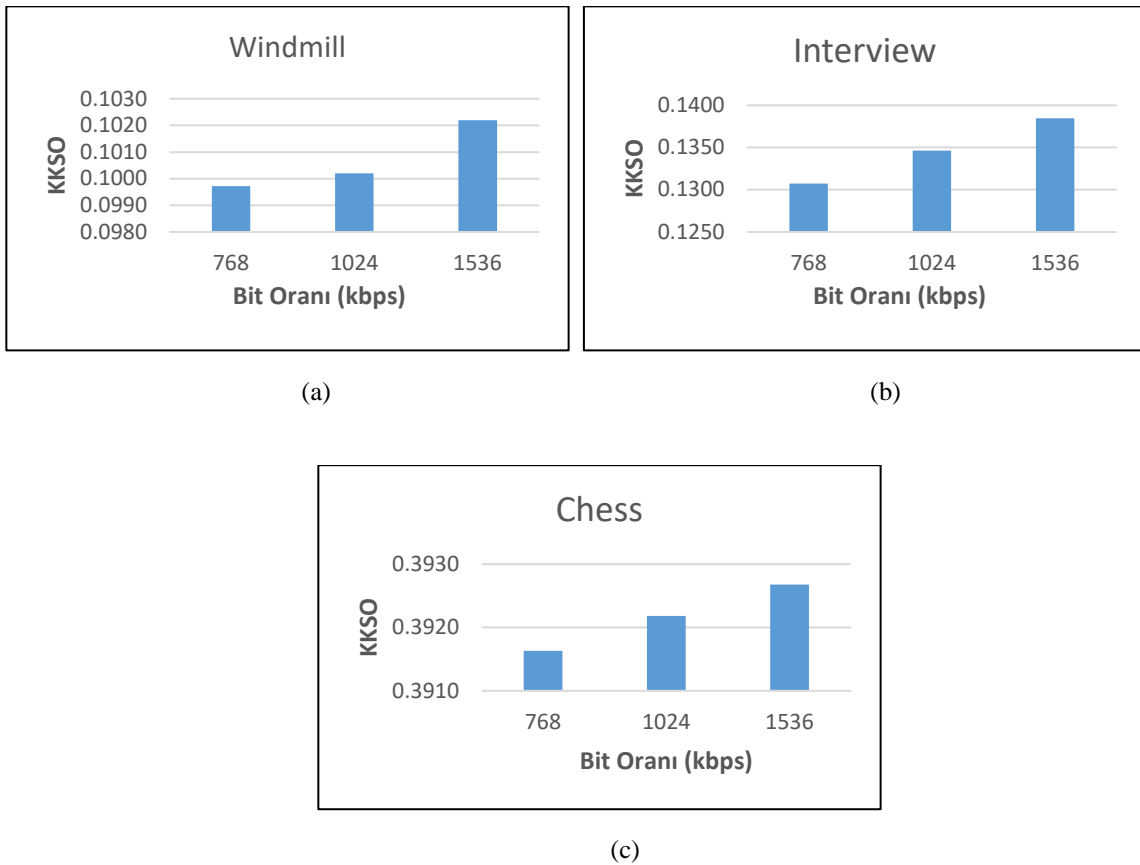
(f)

**Şekil 2:** (a) (b) Chess, (c) (d) Interview ve (e) (f) Windmill videolarının kenar bilgileri çıkarılmış ilk ve son görüntüleri



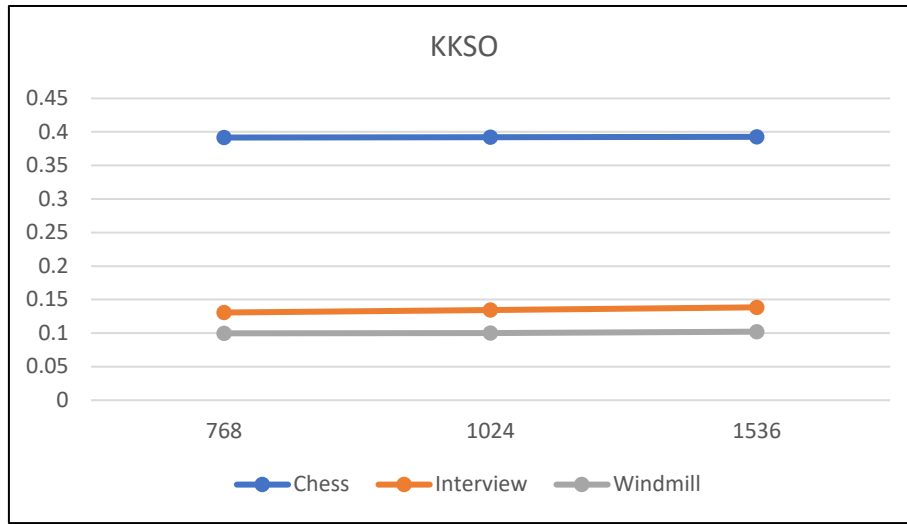


Şekil 3: (a)Chess@768 kbit/s, (b)Chess@1024 kbit/s, (c)Chess@1536 kbit/s, (d)Interview@768 kbit/s, (e)Interview@1024 kbit/s, (f)Interview@1536 kbit/s, (g)Windmill@768 kbit/s, (h)Windmill@1024 kbit/s, (i)Windmill@1536 kbit/s için KKSO eğrileri



Şekil 4: Farklı bit oranlarında kodlanan (a) Windmill, (b) Interview (c) Chess videolarının KKSO değerleri

Şekil 5'te tüm videoların kodlama parametreleri için (768, 1024 ve 1536 kbps) hesaplanmış KKSO değerleri kolay karşılaştırılabilir amaçla bir arada verilmiştir. Şekilden görüleceği üzere, en fazla KKSO değeri tüm kodlama parametreleri için Chess videosuna (0,39) aittir. Interview ve Windmill videoları ise sırasıyla 0,13 ve 0,10 KKSO değerlerine sahiptir. Bu çıkarımlar Şekil 2'deki gözlemlerle paraleldir.



**Şekil 5:**Farklı bit oranlarında kodlanan Windmill, Interview ve Chess videolarının KKSO değerlerinin karşılaştırılması

## 5. VARGILAR VE GELECEKTEKİ ÇALIŞMALAR

Bu çalışmada, 2B+DH temelli 3B videoların 2B kısımlarının VKD'si için YK tabanlı bir referanssız metrik geliştirilmiştir. Metrik kullanılarak elde edilen sonuçlar metriğin İGS'ye hitap etmedeki verimliliğini göstermiştir. Bundan sonraki çalışmalarda daha fazla parametre analiz edilip geliştirilerek İGS'nin algısına daha da yakın referanssız 3B VKD yapan bir metrik geliştirilmesi hedeflenmektedir.

## REFERANSLAR

D.V.S.X. De Silva, G. Nur, E. Ekmekcioglu, and A. Kondo, "QoE of 3D Media Delivery Systems," Media Networks: Architectures, Applications, and Standards, CRC Press Taylor and Francis Group, May 2012.

G. Nur Yilmaz, "A Depth Perception Evaluation Metric for Immersive 3D Video Services," 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video, 7-9 June 2017.

G. Nur Yilmaz, "A No Reference Depth Perception Assessment Metric for 3D Video," Multimedia Tools and Applications, vol. 74, pp. 6937-6950, Sep. 2015.

JSVM 9.13.1. CVS Server [Online]. Available Telnet: garcon.ient.rwth aachen.de:/cvs/jvt

Maria G.Martini, Chaminda T.E.R. Hewage, BarbaraVillarini, "Image Quality Assessment based on Edge Preservation", Signal PProcessing: Image Communication, vol. 27, pp.875-882, September 2012.

Maria G.Martini, Chaminda T.E.R. Hewage, BarbaraVillarini, Image quality assessment based on edge preservation, Brajesh Patel, Neelam Patel, Motion Detection based on multi frame video under Surveillance System, International Journal of Emerging Technology and Advanced Engineering 2 (1), January 2012.

Q. Huynh-Thu, P. Le Callet, and M. Barkowsky, "Video Quality Assessment: From 2D to 3D Challenges and Future Trends," 17th IEEE International Conference on Image Processing, Hong Kong, 26-29 Sep. 2010.

Wei Fu; Xiaodong Gu; Yuanyuan Wang, "Image Quality Assessment Using Edge and Contrast Similarity," IEEE International Joint Conference on Neural Networks, 1-8 June 2008.