

H264/AVC ile Sıkıştırılmış Videoda Sahne Değişimi Kestirimi İçin Bir Yaklaşım

An Approach to Detect Scene Cuts in H264/AVC Coded Video

Gökhan Şimşek, Şevket Gümüştekin

Elektronik ve Haberleşme Mühendisliği Bölümü
İzmir Yüksek Teknoloji Enstitüsü, 35430 İzmir, Türkiye
{gokhansimsek, şevketgumustekin}@iyte.edu.tr

Özetçe

Bu çalışmada ISO/ITU H264/AVC video kodlaması kullanılarak kodlanmış bir video dizisinde makro-öbek bileşenleri çeşitleri analiz edilerek ani ve dereceli sahne kesikleri kestirimi uygulamaları için metodlar önerilmiştir. Sıkıştırılmış sahada yapılan analiz, video dizisinin tamamen sıkıştırmayı çözme sürecine gerek duyulmadan sahne geçişlerinin sezimini gerçekleştirmiştir. Önerilen algoritmayı gerçekleştirmek için çeşitli test video dizileri ISO/ITU H264/AVC kodlanarak QCIF çözünürlüğüne indirgenip, kodlanmıştır ve kısmi sıkıştırmayı çözme işlemi gerçekleştirilerek makro-öbek tipleri istatistikleri incelenmiştir. Video dizisinin kod çözümü sırasında, ISO/ITU H264/AVC standardının ayrıştırma, kısmi kod çözme süreçleri, sözdizim ve anlambilimsel özellikleri kullanılmış ve önerilen metodlar referans yazılım üzerinde uygulanmıştır. Uygulama ISO/ITU H264/AVC kodlanmış sayısal video kütüphanelerinde video özetleme, soyutlama uygulamalarını hedef alan alt seviye araçlar oluşturabilir.

Abstract

In this paper, we provide methods to detect abrupt and gradual scene changes on ISO/ITU H264/AVC coded video sequence using macro-block type information. The analysis is realized in compressed domain without the need of full decompression to pixel domain. In order to realize the proposed methods, several video test sequences are constructed in QCIF resolution. The syntax and semantics, parsing, decoding processes of ISO/ITU H264/AVC is followed to realize such analysis and proposed methods are realized using reference software. The application addresses emerging standard ISO/ITU H264/AVC coded video databases as low level tools to realize a video summarization, abstraction or segmentation system.

1. Giriş

Günümüzde çokluortam video içeriğinin miktarında önemli miktarda artışlar gözlemlenmektedir. Bu içerik DVD videosu, sayısal televizyon, ısmarlama video, video konferans gibi çeşitli ortamlarda kullanıcıya sunulmaktadır. Bu büyük video miktarı, insan emeği olmaksızın arama, özetleme, dizinleme gibi video kütüphaneleri uygulamaları ihtiyacı doğurmaktadır. Bunu gerçekleştirmek için kullanılacak ve anlamlandırılabilir temel bilgi, video dizisindeki sahne geçiş anlarıdır. Sahne kamerasının kayıda başlamasından durdurulmasına kadar kayıt edilen video dizisi şeklinde

tanımlanabilir. Ardışıl sahneler arasında sahne kesiklerini oluşturan sınır durumları vardır.

Sahne geçişlerindeki sınır durumlarını otomatik olarak ayırt etme işlemine sahne kesigi sezimi adı verilir, ve literatürde bu metoda anahtar çerçeve özetlemesi olarak isimlendirildiği de görülebilir [1][2]. Sahne geçişleri video kaydı uygulamasına bağlı olarak çeşitli şekillerde gözlemlenebilir. Bunlardan bazıları ani, dereceli, sönümlenmeli geçişler olabilir [3]. Sahne kesigi sezimi işlemi video ayrıştırma işlemi için temel bir araç olarak kullanılmaktadır.

Literatürde sahne kesiklerini sezme amacıyla çeşitli yordamlar önerilmiştir. Bu yordamlar temelde doğrudan pixel sahasında gerçekleştirilen histogram bilgilerini kullanan, sıkıştırılmış sahada kodlayıcının ürettiği sıkıştırılma parametrelerin özelliklerini kullanan (DCT Dönüşümü Katsayıları, hareket vektörleri) ve metin tanıma bilgilerini kullanan yordamlardır [4].

Piksel sahasında sahne kesik sezimi çerçeveden çerçeveye gerçekleşen renk ve ışıklılık işaretlerinin dağılımına bakılarak yapılır [5]. Piksel sahasında görünen süreksizlik bilgileri de kullanılabilir. Uzay-zamansal video elemanlarının istatistiksel dağılımları belirli bir eşikleme seviyesinden geçirilmektedir. [6]

Sıkıştırılmış sahada gerçekleştirilen işlemler, kodlama parametrelerinin istatistiksel davranışlarını inceler. Bu parametreler kısmi kod-çözme süreci ile elde edildiği için, pixel sahası işlemlerine göre hesaplanabilirlik anlamında etkin çözüm sağlamaktadır. Kullanılabilecek parametreler DCT katsayılarının DC ya da AC katsayıları [1], hareket vektörleri [7], makro-öbek tipleri bilgileri [8] olabilir.

MPEG kodlayıcının özelliklerine bağlı olarak sahne kesiklerini ayırt etmek amacı ile çeşitli öznitelik bulma çeşitleri kullanılır. [9]

Bu çalışmada doğuş sürecini tamamlamak üzere olan ISO/ITU H264/AVC kodlanmış video dizisinde, makro-öbek tipleri özelliklerinin dağılımına dayanılarak sahne kesigi kestirimi gerçekleştirilmiştir. Bu amaçla ISO/ITU H264/AVC [10] kodlayıcının çıktısı ve özellikleri kullanılmıştır.

H264/AVC video kodlaması, ITU-T Video Coding Experts Group (VCEG) ve ISO/IEC Moving Picture Experts Group organizasyonlarının birleşmesinden oluşan Joint Video Team (JVT) tarafından tanımlanmış, uygulama alanı internetten video aktarımından, Yüksek Tanımlı (HD) televizyon alanına

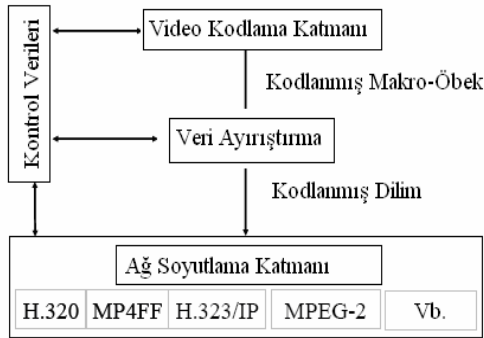
ölçeklendirilebilecek video kodlama standartıdır. Bu uygulamada H264/AVC video dizisi, makro-öbek bilgisi sahasına kadar çözülmüş ve çeşitli sahne kesikleri senaryoları incelenmiştir. Sıkıştırılmış sahada toplanan istatistiksel veriler, video dizisinin içeriği hakkında özetsel bilgiler sunmaktadır.

Bu çalışmanın anlatımı şu şekilde düzenlenmiştir. 2. kısımda H264/AVC kodlamasının uygulamada kullanılan özellikleri tanıtılmakta, 3. kısımda önerilen test düzeneği ve senaryolar tanıtılmakta, 4. kısımda ise sonuçlar ve geliştirilmeye açık öneriler yer almaktadır.

2. H264/AVC Yapısında Sahne Geçişlerinin Etkisi

H264/AVC kodlama sistemi [10] temelde, karma DCT bazlı hareket dengelemesi içeren (MC/DCT) video kodlama yöntemidir. Kodlama verimini arttırmak amacıyla var olan video kodlamalarına çeşitli iyileştirmeler değişiklikler eklenmiştir. [11] H264/AVC ün çeşitli Ağ altyapılarını desteklemek amacı ile video kodlama seviyesi, video paketleri iletim ağı seviyesinden ayrıştırılmıştır.

Bu çalışmada, makro öbek tipleri analizi yapmak amacı ile Şekil 1’de de görülebilecek video kodlama seviyesi üzerinde çalışılmıştır.



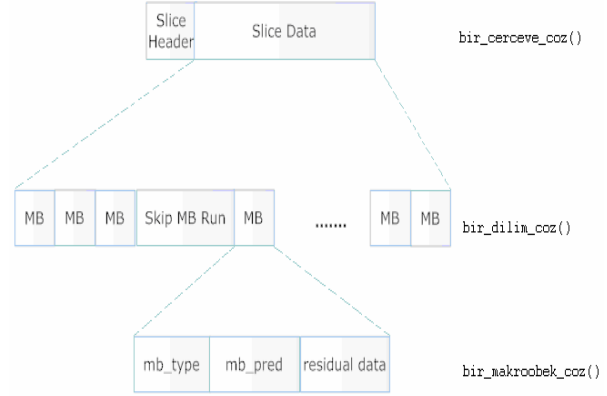
Şekil 1: H264 Elemanları Dağılım Düzlemi

H264/AVC standartında I, P ve B dilim tiplerine ek olarak, SI ve SP dilim tipleri tanımlanmaktadır, bu çalışmada kullanılan ana profilde SI ve SP tipinde dilimler kullanılmamaktadır. I dilim tipi, temel olarak I tipte makro-öbek tiplerini barındırır. P ve B tipte dilimlerde ise I, P tiplerinde makro öbek tipleri, kodlayıcının öngörü kararlarına göre dilimlere yerleştirilirler. [11]

Sahne kesim anlarını ayırtmak amacı ile yapılan makro-blok analizi, I tipi video dilimleri içerisinde bulunan I tipi makro-öbek dağılımını ve P tip dilimlerde bulunan I tipi makro-öbek dağılımı bilgisini analizini kapsar. Bu analizin amacı, kodlayıcının kodlama sırasında zamansal düzlemde yaptığı öngörülerin sahne değişimleri anlarında önce ve sonra çerçevelerden başvuru alamayarak, uzaysal öngörülerini yoğunlukla kullandığı bilgisini ayırt etmektir. Bu da I tipi makro-öbek aktivitesi anlamına gelmektedir.

Bu analizin yapılabilmesi için, kodlanmış bit katarı içerisinde kuşatılmış makro-öbek başlık bilgileri dilim başına kümülatif olarak biriktirilmelidir. (Şekil 2)

Belirlenmiş eşik seviyesini aşan I tipi makro-öbek aktiviteleri potansiyel sahne kesik anlarını belirler. Bu anlar, gösterici sıralaması sayısı (POC) ile eşleştirilmesi durumunda, kesik anlarındaki kodu çözülmüş resim elde edilmiş olur.



Şekil 2: Kodlanmış bit katarının kuşatılmış yapısı

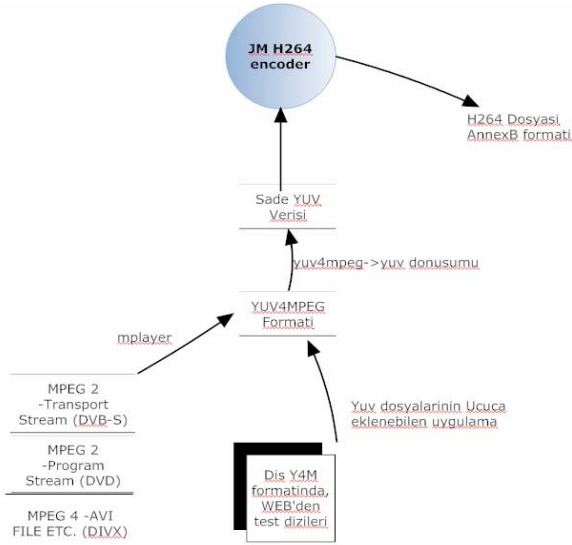
H264/AVC farklı veri taşınması senaryolarında verimli ve hatasız iletim sağlamak amacı ile Esnek Makro-öbek sıralaması (FMO) özelliği sunmaktadır [11]. Bu çalışmada bu özellik kullanılmayıp, tek frame tek resim dilimine gelecek şekilde kodlama yapılmıştır. Bu da makro-öbeklerin dilimler içerisinde satır satır tarama sıralaması ile dizildiğini garantiler, ancak standart bu konuda büyük esneklikler sunmaktadır [10].

Sahne geçişlerinin tek çerçeve geçişi yerine, birden çok çerçeve geçiminde gerçekleşmesi durumunda dereceli ya da sönümlenmeli sahne geçişleri oluşmaktadır. Bu geçiş durumunda I tipi makro-öbek aktivitesi artan ve azalan yokuş örüntüsü izler. Yokuşun başlangıcı sahne geçişinin başlangıcını oluşturur.

Makro-öbek tipi dağılımı sahne kesimleri konusunda olduğu kadar, video dizisindeki yakınlaştırma eylemi konusunda da fikirler sağlamaktadır. Yakınlaştırma, kamera gezdirme, ya da uzaklaştırma eylemi sırasında, video kodlayıcının öngörülerini göreceli olarak az sayıda da olsa, I tipi makro öbekler kullanarak, uzaysal başvuru kararları almasına neden olmaktadır.

2.1. Önerilen Test Altyapısı

H264/AVC kodlanmış video dizileri oluşturmak amacı ile şekil 3’te görülen test altyapısı kurulmuştur. Sistemin girdileri sıkıştırılmamış YUV video verileri, DVB kaynaktan kayıt edilmiş MPEG 2 TS ya da DVD ortamında kayıtlı MPEG2 PS video kaynakları olabilir. Bu girdiler JM referans yazılımı kodlayıcısı ile IPBPBPB.... Çerçeve grupları sıralaması ile kodlanmıştır [12]. Kodlamanın çıktısı “Main profilde”dir. Bu kodlama profili entropi kodlaması olarak CABAC metodlarını kullanmaktadır [10].

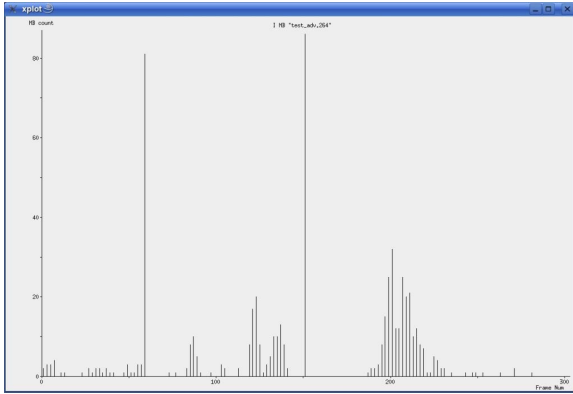


Şekil 3: Önerilen Test Dataseti üretimi akışı

3. Deneysel Sonuçlar

Referans kod çözücü üzerinde makroblok tiplerine göre analiz yapan test yazılımında kullanılmak üzere çeşitli video test dizileri kullanılmıştır.

İlk senaryoda QCIF çözünürlükte Suzie ve Trevor dizileri ardarda eklenerek yapay bir sahne geçişi elde edilerek kodlama yapılmıştır. Şekil 4 te P tip dilimler üzerinde bulunan I tipi makro-öbeklerin çerçeve dizinine göre dağılımı görülmektedir. Trevor dizisi 150 çerçevedir ve çerçeve 151 de Suzie dizisi başlamaktadır.

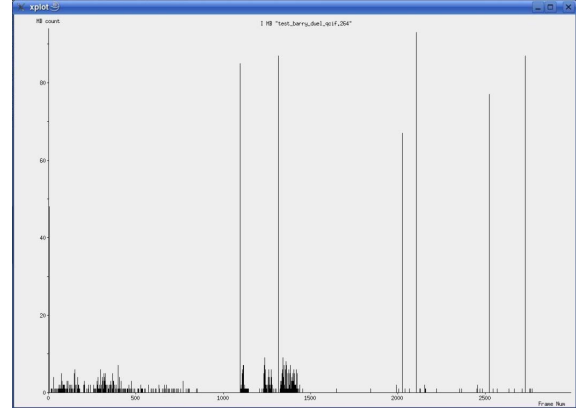


Şekil 4: P tip dilimler üzerinde bulunan I tipi makro-öbeklerin çerçeve dizinine göre dağılımı

Bütünsel olarak belirlenen δ_p eşik değeri toplam makro-öbek sayısının yarısı olarak ani alındığında sahne geçiş anları gözlemlenebilir.

Bir diğer test dizisi, 1975 Barry Lyndon Filminden oluşturulmuştur. Toplam 6 ani sahne geçişi olan bir video dizisidir. P tip dilimler üzerinde bulunan I tipi makro-öbeklerin çerçeve dizinine göre dağılımı Şekil 5 te görülmektedir. Aynı sahnede bulunan sahne kesikleri anlarında çerçeveleri kod çözücü belleğinden alıp, yuv

formatından JPEG formatına çevirip, tek bir JPEG mosaik görüntüsüne çeviren programın çıktısı Şekil 6 dadır.



Şekil 5: P tip dilimler üzerinde bulunan I tipi makro-öbekler



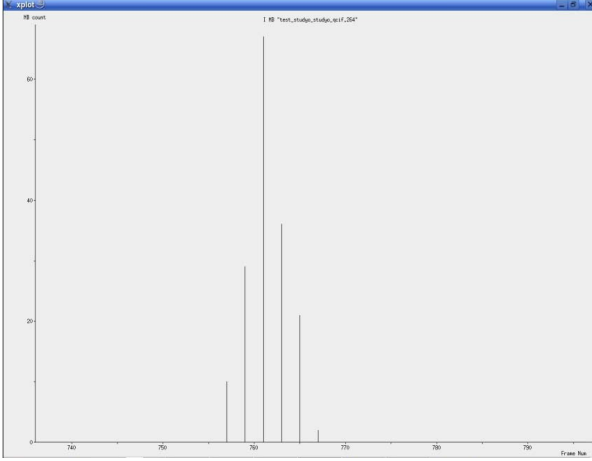
Şekil 6: Sahnenin Özeti (Barry Lyndon Filmi, 1975)

Dereceli, ayrışıklı sahne geçişlerini analiz etmek amacı ile yerel bir televizyon kanalından oluşturulmuş test dizisinden alınan görüntüler Şekil 7'de görülebilir.



Şekil 7: Ayrışıklı Sahne geçişi

Ayrışıklı sahne geçişinde görülen yokuş örüntüsünün yaklaşılmış I makro-öbek sayısı / Çerçeve dizini Şekil 8 de görülmektedir.



Şekil 8: Ayrışıklı sahne geçişinde görülen yokuş örüntüsünün yakınlaştırılmış I makro-öbek sayısı / Çerçeve dizini

4. Sonuçlar:

Bu çalışmada H264/AVC kodlanmış bir video dizisinde, video özetlemeleri, bölütlemeleri, ayırıştırma uygulamalarını hedef almak adına alt seviye araç olabilecek sahne kesimi kestirimi için sıkıştırılmış sahada, makro-öbek istatistiğine dayalı olarak çalışan metodlar tanıtılmıştır. Videonun tamamını kullanıcı tarafından incelemeye gerek bırakmazsınız bir özet sağlayabilecek bu uygulama, makro-öbek tipleri yerine standart tarafından tanımlanmış daha farklı sözdizim öğelerine istatistiksel bir karar süreci izlenerek geliştirilebilecek bir altyapı oluşturmayı hedeflemektedir.

Çalışmanın gelecek aşamalarında şu adımlar atılması hedeflenmektedir.

- Sahne kesikleri arasındaki hareket aktivitesi kullanıcıya özetlenmesi. Bunun için önerilen çözüm hareket vektörlerinin yön ve şiddet istatistiğinin analizidir.
- Alt seviye araçların üst seviyede video anlambilimsel öğelerle birleştirilip video indexleme mekanizmasının tasarımı

5. Kaynakça

- [1] Yeo B. L., and Lui B., "Rapid Scene Analysis on Compressed Video", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol 5, No 6, 1995.
- [2] Gargi U., Rangachar K. and Strayer Susan H, "Performance Characterization of Video-Shot-Change Detection Methods", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol 10, No 1, 1995.
- [3] Mee-Sook L. and Bon-Woo H.; Sanghoon S. and Seong-Whan L, "Automatic Video Parsing Using Shot Boundary Detection and Camera Operation Analysis", *Fourteenth International Conference on Pattern Recognition* Vol. 2, pp. 1481 – 1483, 1998
- [4] Aslandogan Y. A. and Yu Clement T, "Techniques and Systems for Image and Video Retrieval", *IEEE Transactions on Knowledge and Data Engineering*, Vol 11, No:1, 1999.
- [5] Tse K., Wei J., and Panchanathan, "A Scene Change Detection Algorithm for MPEG compressed video sequences.", *Canadian Conference on Electrical and Computer Engineering*, Vol. 2, pp. 827 – 830, 1995

- [6] Hsu, P.R. and Harashima H., "Detecting scene changes and activities in video databases", *IEEE International Conference on Acoustics, Speech and Signal Processing*, Volume 5, pp. V/33- V/36, 1994.
- [7] Mezaris V., Kompatsiaris I., Boulgouris N.V. and Strintzis M.G., "Real-Time Compressed Domain Spatio-Temporal Segmentation and Ontologies for Video Indexing and Retrieval", *IEEE Transactions on Circuits and Systems for Video Technology*, Volume 14, Issue 5, pp 606-621, 2004.
- [8] Pei S. C, and ChouY. Z., "Efficient MPEG Compressed Video Analysis Using Macroblock Type Information", *IEEE Transactions on Multimedia*, Volume 1, Issue 4, pp. 321-333, December, 1999.
- [9] Gerek Ö. N. and Altunbasak Y, "Key frame detection from MPEG video data", *IS&T/SPIE Visual Comm. Image Proc., VCIP'97*, San Jose, CA, vol. 3024, part. 2, pp. 920-925, 12-14 Feb. 1997
- [10] Richardson, I. E. G., *H.264 and MPEG4 Video Compression- Video Coding for Next Generation Multimedia*, Wiley & Sons, England, pp. 123, 2003.
- [11] T. Wiegand, G. Sullivan, A. Luthra, "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Team (ITU-T Rec.H.264 | ISO/IEC 14496-10 AVC)" May, 2003
- [12] Sühring K." *H.264/MPEG-4 AVC Reference Software Manual*", (Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG), pp. 4.1, 2005.