

# Arşiv Filmlerinde Zincirleme Sahne Geçişlerinin Algılanması

## Dissolve Detection for Archive Films

M. Kemal Güllü      Oğuzhan Urhan      Sarp Ertürk

Kocaeli Üniversitesi  
Elektronik ve Haberleşme Mühendisliği Bölümü  
41040 İzmit, Kocaeli  
{kemalg, urhano, sertur}@kou.edu.tr

### Özetçe

Bu çalışmada, arşiv filmlerinde zincirleme sahne değişimlerinin algılanması için piksel temelli özgün bir yöntem önerilmektedir. Önerilen yöntemde, yakınlaştırma-uzaklaştırma etkileri ile film çerçevelerinde oluşan görsel bozukluk etkilerinin azaltılması için ilk olarak video çerçeveleri alt-örneklendirilmiştir. Aşırı değişim maskesinin oluşturulmasıyla yüksek oranda değişim gösteren pikseller değerlendirilmeye alınmayarak, zincirleme sahne geçişlerinde nesne ve kamera hareketlerinin etkisi azaltılmaktadır. Film içerisinde, otuz çerçeve boyunca piksellerin artma, azalma ve değişim özellikleri takip edilmekte ve bu özelliklerin oranı ile oluşturulan metriğin yerel tepe noktaları bulunarak zincirleme sahne değişimleri tespit edilmektedir.

### Abstract

This paper proposes a novel pixel based method for the detection of dissolves in archive films. Initially video frames are sub-sampled in order to reduce effects caused by zoom and visual degradations encountered in the video. Object and camera movement effects are reduced using an excessive change mask that rules out pixels that display extreme variations. Ascending, descending and constant pixel characteristics are tracked within a period of thirty frames and the local peaks of a metric constructed using these characteristics are utilized for dissolve detection.

### 1. Giriş

Son yıllarda, sayısal teknolojiye paralel olarak, sayısal videoların çoklu ortam ve internet uygulamalarında kullanımı artmıştır. Videoda, ani geçişler, dereceli geçişler (çeşitli video etkileri), ışıklılık değişimleri, yakınlaştırma-uzaklaştırma etkisi, nesne ve kamera hareketleri gibi etmenlerden dolayı ardışıl çerçeveler arasında farklılıklar oluşabilmektedir. Ani ve dereceli geçişler şeklinde oluşan sahne değişimlerinin algılanması sayısal videoların düzenlenmesi, indislenmesi ve yeniden elde edilmesi gibi uygulamalar açısından önemlidir. Sahne değişimlerinin algılanmasında en büyük zorluk, diğer etmenlerden dolayı çerçevelerde meydana gelen farklılıklar nedeniyle sahne değişimlerinin ideal özellikte gerçekleşmemesidir.

Ani sahne geçişlerinde bir sahne bitiminden sonra yeni bir sahne başlamaktadır ve bu nedenle geçiş anındaki ilk çerçeve eski sahneye, bir sonraki çerçeve ise yeni sahneye ait

olduğundan bu çerçeveler arasındaki benzerlik ya yoktur ya da oldukça düşüktür. Dereceli sahne geçişlerinde ise, bir sahneden diğer bir sahneye geçiş dereceli olup belirli bir çerçeve süresince gerçekleştirildiğinden, eski sahneye ait görüntü etkisi yavaşça azalırken yeni sahneye ait görüntü yavaşça baskın hale gelmektedir. Bu nedenle dereceli sahne geçişlerinde ardışıl çerçeveler arasında benzerlikler oldukça yüksektir. Dereceli sahne geçişleri dereceli açılmalı (fade in), dereceli kararmalı (fade out), süpürmeli (wipe) ve zincirleme (dissolve) sahne geçişleri olarak çeşitlendirilebilir. Dereceli sahne geçişleri günümüzde pek çok video kameraya görüntü etkisi olarak dahil edilmiş olup, uygulamada en çok kullanılan dereceli sahne geçişi zincirleme sahne geçişidir.

Zincirleme sahne geçişlerinin karmaşık yapılarından dolayı bu sahne geçişlerinin algılanması için önerilen yöntemler halen yeterli başarıyı sağlayamamaktadır. Bu yüzden, zincirleme sahne geçişlerinin algılanmasına yönelik çalışmalar önemini korumaktadır. Zincirleme sahne geçişlerinin bulunması üzerine değişik yöntemler önerilmiştir. Önerilen yöntemler genel olarak histogram temelli yöntemler, özellik temelli yöntemler, piksel temelli yöntemler, model temelli yöntemler ve hareket temelli yöntemler olarak sınıflandırılabilir. [1]'de önerilen histogram temelli yöntem ile, renkli görüntülerde ardışıl çerçeveler arasındaki histogram farklarına bakılarak ani ve dereceli sahne geçişlerinin bulunması amaçlanmıştır. Arşiv filmlerinin genelde siyah-beyaz olduğu göz önünde bulundurulduğunda, siyah-beyaz videolar için farklı sahnelerde benzer histogramlar olabilmemesi, film içerisindeki görsel bozukluklar ile nesne ve kamera hareketinden dolayı histogramda önemli değişimlerin meydana gelebilmesi nedeniyle histogram bilgisini kullanan yöntemler yeterli başarıyı gösterememektedir. Özellik temelli yöntemlere bir örnek [2]'de verilmektedir. Bu yöntemde, ardışıl çerçevelerde giren ve çıkan kenarların istatistiklerine bakılarak zincirleme sahne geçişlerinin algılanması yoluna gidilmiştir. İstatistiklerin değişimi, önceden belirlenmiş sabit zincirleme sahne geçiş örüntüsüne eşlenerek zincirleme sahne geçişleri belirlenmeye çalışılmaktadır, fakat bu yöntem ile elde edilen sonuçlardaki başarıyı da düşüktür. Piksel temelli yöntemlerde, piksel düzeyinde istatistikler kullanılabilir. Kullanılan en temel yöntem, ardışıl imgelerin ortalamalarına ve değişimlerine bakmaktır [3,4]. Bu yöntemlerin en büyük eksikliği, sahne geçişlerinin yanı sıra, video içerisindeki hareket eden nesnelere ve kamera hareketine karşı duyarlı olmalarıdır. [5]'de belirli uzunlukta çerçeveler arası piksel ışıklılık değişimlerinin istatistikleri çıkartılarak, önceden belirlenmiş olan sahne geçiş modellerine ilişkilendirme yapılarak sahne

geçişlerine karar verilmektedir. Bu yöntemin eksikliği ise, yakınlaştırma-uzaklaştırma etkisine, yerel ve kamera hareketine karşı dayanıksız olmasıdır. Hareket temelli yöntemlerde temel olarak çerçeveler arası hareket değişimine bağlı olarak sahne geçişlerinin bulunması yoluna gidilmektedir. [6]'da videodaki her bir çerçeve için hareket değişim oranı ve çerçeveler arası ilişki katsayısı ortalamasını kullanan hareket temelli bir karşılaştırma yöntemi verilmiştir. Çerçeveler arası ilişkinin az olduğu ve hareket değişim oranının yüksek olduğu noktalar, sahne değişiminin sınır noktaları olarak tanımlanmaktadır. [7]'de önerilen yöntemde ise belirli çerçeve sayısı aralığında oluşan zincirleme sahne geçişi modeline bağlı kalınarak, çerçeveler arası artan ve azalan piksel yoğunluklarının sayıları bulunmaktadır. Bu değer, değişken bir eşikle karşılaştırılarak çalışılan aralıktaki değişimin zincirleme sahne geçişi olup olmadığına karar verilmektedir.

Bu bildiriye, video sahneleri arasındaki zincirleme sahne geçişlerinin bulunmasına yönelik piksel temelli yeni bir yöntem önerilmektedir. Önerilen yöntem ile, kamera yakınlaştırma-uzaklaştırma etkisi, yerel ve kamera hareketine karşı daha fazla dayanıklılık sağlanmaya çalışılmıştır.

## 2. Önerilen Yöntem

Bu çalışmada, video sahneleri arasındaki zincirleme sahne geçişlerinin bulunmasına yönelik, piksel temelli yeni bir yöntem önerilmiştir. Önerilen yöntemde öncelikle imgelerin alt-örneklenmesiyle yakınlaştırma-uzaklaştırma etkileri ile film çerçevelerinde oluşan görsel bozuklukların etkisinin azaltılması yoluna gidilmiştir. Daha sonra, video çerçevelerindeki piksel değerlerinin değişimlerinin takip edilmesiyle oluşturulan metrik kullanılarak zincirleme sahne geçişleri algılanmaya çalışılmaktadır.

İdeal bir doğrusal zincirleme sahne geçişinde çıkan ve giren sahneler sabit olarak ele alındığında, oluşturulan zincirleme sahne geçişi sırasındaki bir çerçeve ( $Z_t$ ), giren ( $G$ ) ve çıkan

( $C$ ) sahnelerin doğrusal birleşimi şeklinde denklem (1)'de gösterildiği gibi tanımlanmaktadır [7].

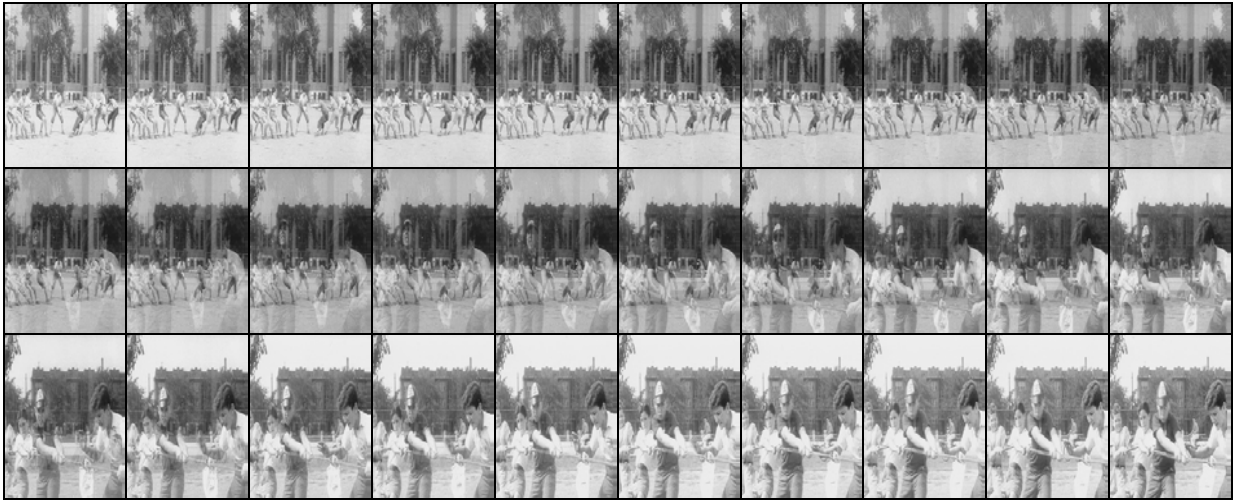
$$Z_t(x, y) = \frac{t}{n}G(x, y) + \left(1 - \frac{t}{n}\right)C(x, y) \quad \forall t \in [0, n] \quad (1)$$

Burada  $Z_t(x, y)$ ,  $t$ . çerçeve için  $(x, y)$  noktasındaki piksel yoğunluğunu,  $n$  zincirleme sahne geçişi uzunluğunu göstermektedir. Örneğin; otuz çerçevelik bir zincirleme sahne geçişinde giren çerçevenin  $t=1$  anında  $Z_t$ 'ye katkısı  $1/30$  iken çıkan çerçevenin katkısı  $29/30$ 'dur. Pratikte ise nesne ve kamera hareketleri gibi etmenlerden dolayı ideal bir zincirleme sahne geçişine nadiren rastlanmaktadır. Şekil 1'de örnek "Allusa" görüntü dizininden doğrusal bir zincirlemeli sahne geçişi görülmektedir.

İdeal bir doğrusal zincirleme sahne geçişi süresince her bir pikselin yoğunluğunun doğrusal biçimde daima artması, daima azalması veya hiç değişmemesi gerekmektedir. Gerçekte ise bu durumu değiştiren kamera yakınlaştırma-uzaklaştırma etkisi, filmdeki fiziksel bozuklukların görüntüde oluşturduğu düzensizlikler, kamera hareketleri ve yerel hareketler gibi etkenler vardır.

Piksel temelli yöntemlerde, bu etkenlerden dolayı zincirleme sahne geçişinde ardışıl çerçevelerdeki mutlak piksel değişimi çok yüksek değerlere ulaşabilmektedir. Bu da zincirleme sahne geçişlerinin bulunmasını zorlaştırmaktadır. Bu çalışmada önerilen yöntem ile bu problemin üstesinden gelebilmek için, ardışıl çerçevelerdeki piksel değerlerinin mutlak değişimi belli bir değer ile eşiklenerek, çok fazla değişim gösteren ve bu nedenle zincirleme sahne geçişi özelliği taşımayan pikseller aşağıdaki gibi ayırt edilmektedir.

$$M_t(x, y) = \begin{cases} 1, & |I_t(x, y) - I_{t-1}(x, y)| < th_m \\ 0, & \text{diğer} \end{cases} \quad (2)$$



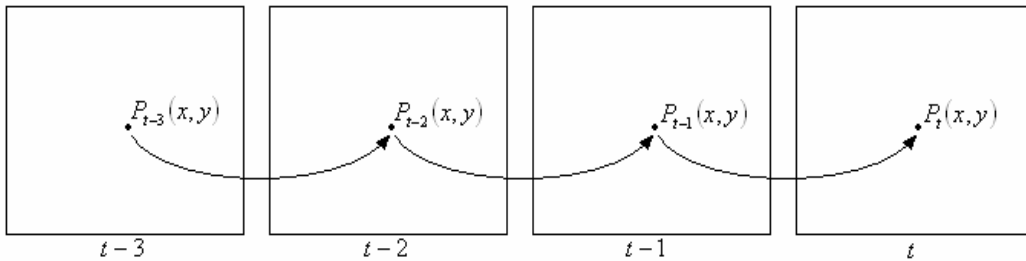
Şekil 1: "Allusa" görüntü dizininden 1126-1156 çerçeveleri arası otuz çerçevelik bir zincirleme sahne geçişi.

Burada  $I_t$  ve  $I_{t-1}$ , sırasıyla video dizinindeki  $t$  ve  $(t-1)$  anlarındaki ardışıl imge çerçevelerini göstermektedir. Oluşturulan  $M_t$  maskesinde zincirleme sahne geçişini algılamada kullanılacak piksellerin bilgisi tutulmaktadır. Zincirleme sahne değişimlerinin algılanması için sadece  $M_t$  maskesinde “1” değeri alan piksellerin kullanılmasıyla kamera yakınlaştırma-uzaklaştırma etkisi, filmdeki fiziksel bozuklukların görüntüde oluşturduğu düzensizlikler, kamera hareketleri ve yerel hareketlerin etkileri büyük ölçüde giderilmektedir. Burada kullanılan eşik değeri, en kısa zincirleme sahne geçişi uzunluğunun otuz çerçeve ve 8-bit/piksel kullanan gri-tonlu imgelerde zincirleme geçişlerde çerçevelerdeki piksel değeri farkının en fazla 255 olabileceği göz önüne alınarak,  $th_m = 255/30$  alınmıştır.

Zincirleme sahne geçişinin yakalanabilmesi için otuz çerçevelik izleme pencereleri ile piksel değerlerinin değişim bilgilerine bakılmaktadır. Zincirleme sahne geçişi boyunca başarıyla artan, azalan ve değişmeyen piksellerin belirlenmesi için artan, azalan ve değişmeyen maskeleri kullanılmaktadır. Şekil 2’de gösterildiği gibi arka arkaya üç kez aynı değişim özelliği gösteren (artan, azalan, veya değişmeyen) pikseller, ilgili maskede 1, diğer maskelerde 0 değerini almakta, arka arkaya üç kez aynı özelliği göstermeyen pikseller ise her üç maskede de 0 değerini almaktadır. Bu üç maske otuz çerçeve boyunca izlenerek aşağıda tanımlanan metriğe göre zincirleme sahne geçişi için aday bölgelere karar verilmektedir.

$$S_t = \max \left( \sum_{i=0}^{w-1} \sum_{j=0}^{h-1} \frac{art}{az + ayn}, \sum_{i=0}^{w-1} \sum_{j=0}^{h-1} \frac{az}{art + ayn} \right) \quad (3)$$

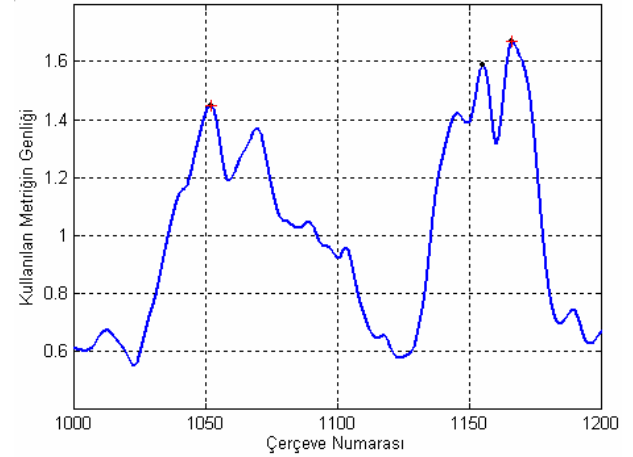
Burada  $(w, h)$  çerçeve boyutlarını;  $art$ , otuz çerçeve boyunca artan piksellerin toplamı (artan maskesinin otuz çerçeve boyunca toplamını);  $az$ , otuz çerçeve boyunca azalan piksellerin toplamı (azalan maskesinin otuz çerçeve boyunca toplamını) ve  $ayn$ , otuz çerçeve boyunca aynı kalan piksellerin toplamını (değişmeyen maskesinin otuz çerçeve boyunca toplamını) göstermektedir.



Şekil 2: Arka arkaya üç kez aynı değişim özelliği gösteren piksellerin tespit edilmesi.

Zincirleme sahne geçişi özelliği gösteren bölgelerde çerçeve piksel değerlerinin bir sahneden diğer sahneye geçerken sahne özelliklerine bağlı olarak düzenli bir artış veya azalış göstermesi beklendiğinden  $S_t$  metriği yüksek bir değer olacaktır. Bu nedenle dizin boyunca  $S_t$  metriğinin takip edilmesiyle zincirleme sahne değişimlerine karar verilebilmektedir. Bu amaçla dizin boyunca sürekli olarak hesaplanan  $S_t$  metriği ani değişimlerin yumuşatılması için ortalama süzgecinden geçirildikten sonra bir eşik değeri ile karşılaştırılmaktadır.

Eşik değerinin üzerine çıkan aday bölgelerde yerel tepe noktaları tespit edilerek zincirleme sahne değişimine karar verilmektedir. Kullanılan metrik nedeniyle aynı zincirleme sahne geçişi içerisinde birden fazla yerel tepe nokta oluşabildiğinden, birbirine otuz çerçeveden daha yakın olan yerel tepe noktalarından sadece genişliği en yüksek olan değerlendirilmeye alınmaktadır. Şekil 3’de “Allusa” dizini için 1000-1200 çerçeveleri arası  $S_t$  metriğinin değişimi ve tespit edilen iki adet zincirleme sahne geçişi bölgeleri gösterilmektedir.



Şekil 3: “Allusa” görüntü dizininin 1000-1200 çerçeveleri arası bulunan zincirleme sahne geçişleri

### 3. Deneysel Sonuçlar

Sahne değişimi algılama yöntemlerinin performansını değerlendirmek için genellikle kullanılan iki metrik

$$\text{Hatırlama} = \frac{D}{D+K} \text{ ve Kesinlik} = \frac{D}{D+Y} \quad (5)$$

olarak tanımlanmaktadır [8]. Burada D; toplam doğru algılanan zincirleme sahne değişim sayısını, Y; toplam yanlış algılanan zincirleme sahne değişim sayısı, K ise toplam kaçırılan (yakalanamayan) zincirleme sahne değişim sayısını göstermektedir. Tablo1'de 50 adet zincirleme sahne geçişi içeren "Allusa" görüntü dizini için önerilen yöntemin sonuçları ile [7]'de önerilen yöntemin sabit eşikle elde edilmiş sonuçları verilmektedir.

Tablo 1. "Allusa" görüntü dizini için önerilen yöntem ile [7]'de önerilen yöntemin başarımlarının karşılaştırması.

Yöntem	D	Y	K	Hatırlama	Kesinlik
<b>Önerilen Yöntem</b>	46	13	4	<b>0.92</b>	<b>0.78</b>
[7]'de Önerilen Yöntem	45	32	5	0.9	0.59

Başarım sonuçları incelendiğinde önerilen yöntemin, [7]'de önerilen yöntemin sabit eşikle elde edilmiş sonuçlardan biraz daha iyi bir hatırlama oranı verdiği fakat kesinlik oranlarına bakıldığında önerilen yöntemin çok daha iyi sonuç verdiği görülmektedir. Bunun nedeni, önerilen yöntemin kamera yakınlaştırma-uzaklaştırma etkisi, yerel ve kamera hareketi gibi etkenlere karşı daha dayanıklı olmasıdır.

### 4. Sonuç

Bu çalışmada, zincirleme sahne değişimlerinin algılanması için özgün bir yöntem önerilmiştir. Önerilen yöntemde, video çerçeveleri ilk aşamada alt-örnekleyerek yakınlaştırma-uzaklaştırma etkileri ile film çerçevelerinde oluşan görsel bozukluk etkileri azaltılmaktadır. Aşırı değişim gösteren piksellerin değerlendirmeye alınmaması ile nesne ve kamera hareketlerinin bulunduğu zincirleme sahne geçişlerinde başarımlar artırılmıştır. Film içerisinde, otuz çerçeve boyunca piksellerin artma, azalma ve değişmeme özellikleri takip edilerek, önerilen metriğin yerel tepe noktalarının bulunmasıyla zincirleme sahne değişimleri tespit edilmektedir.

### 5. Teşekkür

Bu çalışma, Türkiye Bilimsel ve Teknik Araştırma Kurumu, TÜBİTAK tarafından EEEAG/103E007 nolu araştırma projesi kapsamında desteklenmiştir.

### 6. Kaynakça

[1] Zhang H. J., Kankanhalli A., and Smoliar S. W., "Automatic Partitioning of Full-motion Video", *ACM Multimedia Systems*, pp.10-28, 1993.

- [2] Zabih R., Miller J., and Mai K., "A Feature-Based Algorithm for Detecting and Classifying Production Effects", *Multimedia Systems*, Vol.7, pp.119-128, 1999.
- [3] Truong B.-T., Venkatesh S. and, Dorai C., "New Enhancements to Cut, Fade, and Dissolve Detection Process in Video Segmentation", *Proc. ACM Int. Conf. Multimedia*, pp.219-227, 2000.
- [4] Fernando W., Canagarajah C., and Bull D., "Fade and Dissolve Detection in Uncompressed and Compressed Video Sequences", *Proc. ICIP*, Vol.3, pp.299-303, 1999.
- [5] Yu H., Bozdagi G., and Harrington S., "Feature-based Hierarchical Video Segmentation", *Int. Conf. Image Proc.(ICIP'97)*, pp.498-503, 1997.
- [6] Akutsu A., Tonomura Y., Hashimoto H., and Ohba Y., "Video Indexing Using Motion Vectors", *SPIE Visual Comm. And Image Processing'92 Conf. Proc.*, Vol.1818, pp.1522-1530, 1992.
- [7] Su C. W., Tyan H. R., Mark Liao H. Y., and Chen L. H., "A Motion-tolerant Dissolve Detection", *Proc. IEEE Int. Conf. Multimedia and Expo*, pp.225-228, 2002.
- [8] Li W.-K., and Lai S.-H., "Integrated Video Shot Segmentation Algorithm", *SPIE Conf. on Storage and Retrieval for Media Databases*, pp. 264-271, Santa Clara, California, Jan. 2003.