

# ALT PİKSEL DOĞRULUKTA HIZLI HAREKET KESTİRİMİ İÇİN BİR-BİT DÖNÜŞÜMÜNÜN KULLANILMASI

## FAST SUB-PIXEL MOTION ESTIMATION VIA ONE-BIT TRANSFORM

*Orhan Akbulut, Oğuzhan Urhan, Sarp Ertürk*

Elektronik ve Haberleşme Mühendisliği Bölümü  
Veziroğlu Yerleşkesi, Kocaeli Üniversitesi, 41040, İzmit/Kocaeli

orhanakbulut@gmail.com {urhano, sertur}@kou.edu.tr

### Özetçe

Bu çalışmada bir-bit dönüşümü yoluyla piksel altı doğrulukta hareket kestirimi önerilmiştir. Literatürde hızlı hareket kestirimi için düşük bit gösterimli dönüşümler kullanıldığı halde [1,7] bu dönüşümlerin piksel altı doğruluk için kullanımı konusunda bir çalışma mevcut olmayıp ilk defa bu çalışma kapsamında ele alınmıştır. Bir-bit (1BT) [1] dönüşümü kullanılarak gerçekleştirilen hareket kestirimi  $8 \times 8$  blok boyutu için kötü sonuçlar verebilmektedir. 1BT dönüşümüne ek olarak imgelerin yarım ve çeyrek piksel doğrulukta hareket vektörleri hesaplanabilir. Bu çalışmada yarım ve çeyrek piksel doğrulukla hesaplanan hareket vektörleri ile 1BT dönüşümüne göre daha yüksek başarımlar elde edilmiştir.

### Abstract

In this study sub-pixel motion estimation via one-bit transform is proposed. Low-bit resolution representations are exist in literature for motion estimation [1,7]. However sub-pixel motion estimation employing these kinds of transforms has not tried until now. One-bit (1BT) transform based motion estimation [1] can give inaccurate result for block size of  $8 \times 8$ . Half and quarter pixel motion vectors can be calculated using one-bit transform. In this work higher performance is achieved computing half and quarter pixel motion vectors utilizing one-bit transform than pure one-bit motion estimation.

### 1. Giriş

Video sıkıştırma, video verisinin daha düşük bit oranlarında ifadesi, düşük bellek ve bant genişliği kullanımı için gereklidir. Bilindiği gibi hareket kestirimi ve dengelemesi imge çerçeveleri arasındaki zamansal artıklığı kullanarak sıkıştırma performansını arttırmak amacı ile kullanılmaktadır. İyi bir hareket kestirimi sayesinde etkili bir hareket dengelemesi yapılacak bu sayede verimli bir video sıkıştırma gerçekleştirilecektir.

Hareket kestirimi için video kodlamada genellikle blok uyumlama (block matching) yaklaşımı kullanılmaktadır. Blok uyumlama yönteminde, imge çerçevesi bloklara bölünmekte sonrasında her bir blok bir önceki çerçevede önceden belirlenmiş bir arama penceresi içinde en iyi uyumu verecek şekilde aranmaktadır. En iyi uyumu veren blok aranan blokta farkı alınır ve bu fark bilgisi ile hareket vektörleri kodlanır [2].

Blok uyumlama temelli hareket kestirimi için en iyi sonucu veren arama yaklaşımı tam aramadır. Bu yöntemde aranacak blok, arama penceresi içinde olası bütün noktalarda aranmaktadır. Bu yöntemin dezavantajı ise hesapsal yükün fazla olmasıdır. Daha düşük hesapsal karmaşıklığa sahip olan hareket kestirim yöntemlerinin kullanılması durumunda bile video kodlama işleminin neredeyse yarı zamanı hareket kestirimi işlemi ile harcanmaktadır.

Tam arama yönteminin hesapsal yükünün fazla olmasından dolayı daha düşük işlem yüküne sahip birçok hareket kestirim yöntemi önerilmiştir. Ancak bununla birlikte genel olarak yöntemlerin başarımları da büyük ölçüde düşmektedir. Bu yöntemler genellikle ya arama noktalarının sayısını azaltmakta ya da düşük bit gösterimi ile işlem yükünü düşürmeye çalışmaktadır. Hızlı arama yöntemlerinden bazıları üç adımlı arama (3SS) [3], 2D logaritmik arama (2DLOG) [4], yeni üç adımlı aramadır (N3SS) [5].

Hareket kestiriminde blok uyumlama için çeşitli ölçütler kullanılmaktadır. Diğer bir ifade ile farklı arama stratejileri olduğu gibi farklı blok eşleme ölçütleri de vardır. Minimum hata kriterlerinden biri olan ortalama mutlak fark (MAD - Mean Minimum Absolute Difference) gerçekleştirilmesi kolay olduğu için hareket kestiriminde yaygın olarak kullanılmaktadır. Diğer bir uyumlama kriteri ise ortalama karesel hatadır (MSE - Mean Squared Error). En büyük işaret gürültü oranı (PSNR - Peak Signal to Noise Ratio) bakımından en iyi sonuç bu uyumlama ölçütü kullanıldığından elde edilmektedir. Ancak bu işlemin de donanımsal olarak gerçekleştirilmesinin zor olması donanıma uygun hata ölçütlerinin geliştirilmesine yol açmıştır.

Uyumlama ölçütlerinin daha hızlı hesaplanabilmesi için düşük bit gösterimli hareket kestirim yöntemleri önerilmektedir [1,2,7]. Böylelikle imge çerçevelerinde yapılacak arama işleminde uyumlama ölçütünün daha hızlı hesaplanması sağlanabilir. Bu amaçla kullanılan genel yaklaşım donanımsal olarak gerçekleştirilmesi oldukça basit ve hızlı olan özel veya (XOR) işleminin uyumlama ölçütü hesabında kullanılmasıdır.

Düşük bit çözünürlüğünde hareket kestirimi öneren yöntemlerden birisi bir-bit (1BT) dönüşümüdür. 1BT dönüşümünün temel amacı, imgeleri ikili düzeyde göstererek piksel değerlerinin bit çözünürlüğünü azaltmaktır. Bu yaklaşımda 8-bit çözünürlüğündeki gri tonlu imge çerçevesi (1)'de verilen çekirdek (kernel) ifadesi kullanılarak süzgeçlenmektedir. Bu çekirdek altında çoklu bant geçiren bir süzgeç işlevi görmektedir.

$$K_{i,j} = \begin{cases} 1/25, & i, j \in [0,4,8,12,16] \\ 0, & \text{aksi halde} \end{cases} \quad (1)$$

İmge çerçevesi bu çekirdekten geçirildikten sonra (2)'de verilen ifade ile sadece "0" ve "1" seviyelerden oluşan ikili imgeye dönüştürülmektedir.

$$G_{i,j} = \begin{cases} 1, & F_{i,j} \geq F'_{i,j} \\ 0, & \text{aksi halde} \end{cases} \quad (2)$$

$F'$  çerçevesi, çekirdek matrisi ile  $F$  imge çerçevesinin konvolüsyonu ile elde edilmiştir. 1BT yönteminde blok uyumlama ölçütü hesabında XOR işlemini kullanılmaktadır. (3)'de  $(u, v)$  bloğu için uyumlama ölçütünün hesabı verilmektedir.

$$\|u, v\|_1 = \sum_1^N \sum_1^N u_{i,j} \otimes v_{i,j} \quad (3)$$

Bu eşitlikte,  $\otimes$ ; XOR işlevini göstermektedir.  $u_{i,j}$  ve  $v_{i,j}$  sırasıyla  $u$  ve  $v$  blokların piksel değerleridir. Bloklar ise  $N \times N$  pikselden oluşmaktadır. Bir-bit dönüşümünün blok karşılaştırılmasını XOR işlevi ile gerçekleştirmesinden dolayı, donanımsal olarak bu işlemi gerçekleştirmek hız açısından önemli bir performans artışı sağlamaktadır.

MPEG1, MPEG2 ve H.261 video sıkıştırma standartlarında genellikle  $16 \times 16$  boyutunda bloklar kullanılmakta, ancak tanımlanmış standart bir hareket kestirim metodu bulunmaktadır [6]. [7]'de belirtildiği gibi  $16 \times 16$  blok boyutları için 1BT hız ve kalite açısından iyi sonuçlar üretebilmektedir. Bundan dolayı 1BT, MPEG1, MPEG2 ve H.261 video sıkıştırma standartlarında kullanılabilir. Buna rağmen 1BT,  $8 \times 8$  ve daha düşük piksel boyutlarında bloklar için uyumlama hesabı gerektiren H.264 gibi video sıkıştırma standartlarında düşük blok boyutlarında daha kötü sonuç vermesi nedeni ile etkin sıkıştırmaya olanak sağlayamayacağı için kullanılması mümkün değildir.

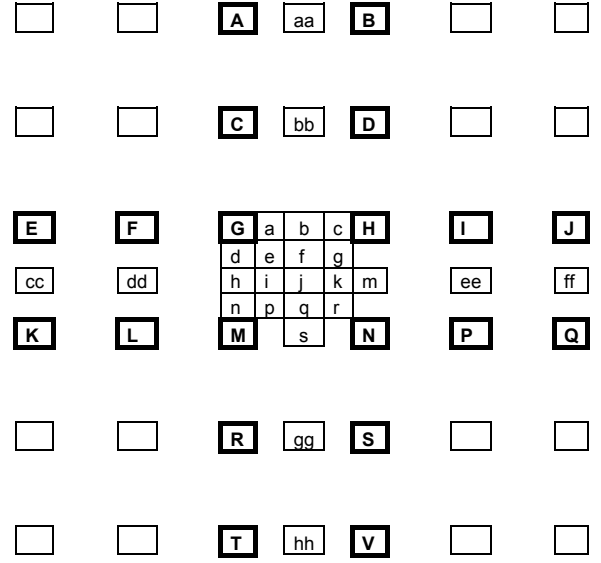
Bu çalışmada, imge çerçevelerinin ara değerlemesi ile elde edilen imge çerçevelerin yine ikili bit uzayında değerlendirilerek hareket kestiriminin özellikle küçük blok boyutlarda iyi sonuç vermesi amaçlanmıştır.

## 2. Piksel Altı Doğruluk Yaklaşımı

Hareket kestiriminde hareket vektörleri genellikle tamsayı piksel değerleri ile sınırlanmaktadır. En iyi blok uyumu için hareket vektörü çözünürlüğü bir pikselin altında bir seviyede olmalıdır. Bu yaklaşımdan yola çıkarak hareket kestirimi performansını arttırmak için imge çerçevelerinin üst örnekleme ile daha yüksek çözünürlükte hareket vektörleri elde edilebilir.

MPEG2 ve H.263 standartlarında yarım (0,5) piksel, MPEG4 ve H.264 standartlarında ise çeyrek (0,25) piksel çözünürlüklü hareket vektörleri hesaplanmaktadır. Daha yüksek çözünürlüklü hareket vektörleri ile daha iyi hareket tahmini elde etmek mümkündür.

Bu çalışmada 0.5 ve 0.25 piksel çözünürlüklü hareket vektörlerini bulmak için, katsayı değerleri (1,-5,20,20,-5,1) olan 6. dereceden FIR süzgeç yöntemi kullanılmıştır. Ara değerlendirme işlemi bu süzgeç kullanılarak gerçekleştirilmiştir.



Şekil 1: FIR süzgeç ile elde edilmiş ara değerlendirme işlemi.

Literatürde çift doğrusal ara değerlendirme süzgeci, Wiener süzgeci ve Telenor süzgeci [8] gibi farklı ara değerlendirme süzgeçleri bulunmaktadır. Bu süzgeçlerde kullanılan katsayı değerleri ve süzgeç uzunluğu, kullanılacak süzgeç yapısına göre değişmektedir. Bu çalışmada Telenor süzgeci kullanılmıştır. 0.25 piksel çözünürlüklü ara değerlendirme işleminin gerçekleştirilmesi iki aşamadan oluşmaktadır. İlk aşama 0.5 piksel çözünürlüklü ara değerlendirme işlemidir. Diğer bir deyişle çerçevenin 2 kat üst örnekleme işlemidir. Sonrasında aynı işlem 2 kat üst örnekleme imgeye uygulanarak imge 4 kat üst örnekleme ile elde edilir.

Şekil-1'de görülen ara değerlendirme ile elde edilmiş yeni piksel değerlerinin hesaplanmasında, Telenor süzgecinin katsayıları kullanılmıştır. Her yeni piksel değeri aşağıdaki eşitliklerle hesaplanmıştır.

- b pikseli gibi 0.5 çözünürlüklü piksel değerlerini hesaplamak için, ilgili pikselin yatay yöndeki kendisine en yakın tamsayı piksel değerleri 6. dereceden süzgeç katsayıları ile çarpılır. Yine aynı şekilde h pikselin değerini hesaplamak için, ilgili pikselin dikey yöndeki kendisine en yakın tamsayı piksel değerleri 6. dereceden süzgeç katsayıları ile çarpılır.

$$b = (E-5F+20G+20H-5I+J) \\ b = ((b+16)/32) \quad (4)$$

$$h = (A-5C+20G+20M-5R+T) \\ h = ((h+16)/32) \quad (5)$$

- j pikseli gibi 0.5 piksel çözünürlüklü piksel değerlerini hesaplamak için ilgili pikselin yatay veya dikey yöndeki kendisine en yakın tamsayı piksel değerleri 6. dereceden süzgeç katsayıları ile çarpılır.

$$\begin{aligned} j &= cc - 5dd + 20h + 20m - 5ee + ff \\ j &= aa - 5bb + 20b + 20s - 5gg + hh \end{aligned} \quad (6)$$

$$j = ((j + 512) / 1024) \quad (7)$$

- a,c,d,n,f,i,q,k pikselleri gibi 0.25 piksel çözünürlüğünde olan yeni piksel değerleri, kendisine en yakın tamsayı piksel değeri ile 0.5 piksel çözünürlüklü piksel değerinin ortalaması alınarak hesaplanır.

$$\begin{aligned} a &= ((G+b)/2), \quad c = ((H+b)/2) \\ d &= ((G+h)/2), \quad n = ((M+h)/2) \\ k &= ((j+m)/2), \quad q = ((j+s)/2) \\ f &= ((b+j)/2), \quad i = ((h+j)/2) \end{aligned} \quad (8)$$

- e,g,p,r pikselleri gibi 0.25 piksel çözünürlüğünde olan yeni piksel değerleri, çaprazında bulunan kendisine en yakın 0.5 piksel çözünürlüklü iki piksel değerinin ortalaması alınarak hesaplanır.

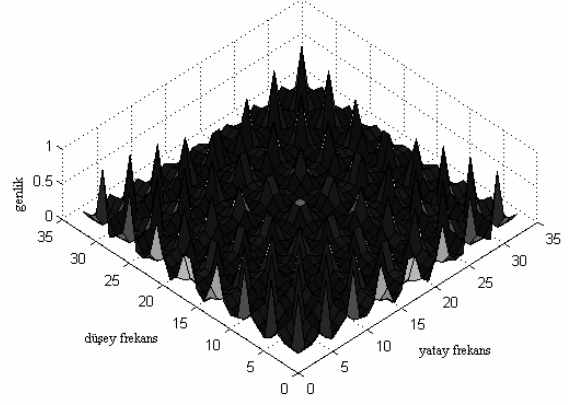
$$\begin{aligned} e &= ((b+h)/2), \quad g = ((b+m)/2) \\ p &= ((h+s)/2), \quad r = ((m+s)/2) \end{aligned} \quad (9)$$

Hareket kestirimi için öncelikle ara değerlendirme işlemi yapılmıştır. Ara değerlendirme işlemi sadece bir önceki çerçeveye uygulanmıştır. Her iki çerçeve de çekirdek süzgeci ile [1]'deki gibi ikili bit uzayına dönüştürülmüştür. Ancak bu aşamada farklı çekirdekler kullanılmıştır. Daha sonra çerçeve içinde referans olan blok, ara değerlendirme ile elde edilmiş bir önceki çerçevede tam arama yöntemi ile XOR işlemi kullanılan uyumlama ölçütü ile aranmıştır. En uygun blok uyumu bulunup hareket vektör elde edilmiştir.

Dikkat edilmesi gereken bir nokta, arama penceresinin boyutunu belirlemede kullanılan maksimum yer değiştirme vektörünün yarım piksel çözünürlük kullanılacaksa 2 katına çıkması, çeyrek piksel kullanılacaksa 4 katına çıkmasıdır. Bir başka önemli nokta ise uygulamada kullanılan çoklu bant geçiren süzgecin boyutudur. 1BT'de kullanılan süzgeç ara değerlendirme ile elde edilen çerçeve üzerine uygulanmamıştır. Çekirdek süzgeci boyutunun ara değerlendirme yöntemi ile artırılması gerekmektedir. Bu amaçla 0.5 ve 0.25 piksel doğruluğu için aynı işlevi görecektir iki farklı çekirdek süzgeci kullanılmıştır. Uygulamada kullanılan çekirdek süzgeçleri sırasıyla eşitlik (10) ve (11)'de verilmiştir. 0.5 piksel doğruluklu hareket vektörünü bulmak için kullanılan çekirdek süzgecinin frekans yanıtı ise Şekil-2'de gösterilmiştir.

$$K_{i,j} = \begin{cases} 1/25, & i, j \in [0, 8, 16, 24, 32] \\ 0, & \text{aksi halde} \end{cases} \quad (10)$$

$$K_{i,j} = \begin{cases} 1/25, & i, j \in [0, 16, 32, 48, 64] \\ 0, & \text{aksi halde} \end{cases} \quad (11)$$



Şekil 2: 0,5 piksel doğrulukta hareket kestirimi için kullanılan çekirdek süzgecinin frekans yanıtı.

### 3. Deneysel Sonuçlar

Önerilen alt piksel doğrulukta bir-bit dönüşümü kullanan hareket kestirimi yaklaşımının başarımını değerlendirmek için birçok deney yapılmıştır.

Şekil-3'de ise bu yaklaşımların nesnel başarımlarını değerlendirmek üzere "Foreman" dizisi için PSNR ölçütünün imge çerçevelerine göre değişimi verilmiştir. Bu sonuçlardan, bu çalışmada önerilen 1BT yönteminin piksel altı doğrulukta kullanılmasını öneren hareket kestirimine yaklaşımının, standart 1BT hareket kestirimine göre daha yüksek başarımlar gösterdiği açıkça görülmektedir.

Tablo 1 ve Tablo 2'de "Football", "Flowergarden", "Mobile", "Tennis", "Coastguard", "Foreman" dizileri için 16×16 piksel ve 8×8 piksel boyutunda bloklar kullanılarak yapılan farklı hareket kestirimi yaklaşımlarının PSNR sonuçları verilmektedir. Bu tablolarda MAD uyumlama yaklaşımının yanı sıra MAD uyumlama yaklaşımının alt piksel doğrulukta verdiği sonuçlar da gösterilmektedir. Tablolardan açıkça görüldüğü gibi alt piksel doğrulukta yapılan hareket kestirimi hem MAD hem de 1BT yaklaşımları için daha yüksek başarımlar sağlamaktadır. Hatta, çeyrek piksel doğrulukta kullanılan 1BT yaklaşımı, standart 1BT yaklaşımına göre bazı diziler için 1.5dB'den daha yüksek başarımlar sağlamaktadır. Yine çeyrek piksel doğruluk için 1BT yaklaşımı, 2BT yaklaşımından genel olarak daha yüksek bir başarımlar göstermektedir.

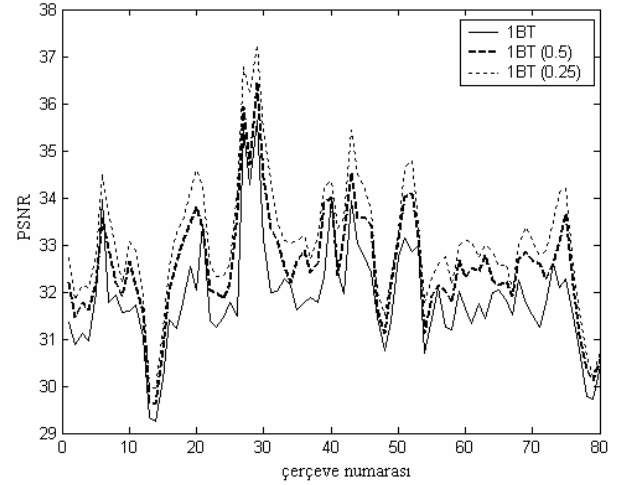
### 4. Sonuçlar

Hızlı hareket kestirimi için bir-bit dönüşümü dâhil düşük bit gösterimli dönüşümler literatürde bulunduğu halde bu dönüşümlerin piksel altı doğruluk için kullanımı konusunda bir çalışma mevcut olmayıp ilk defa bu çalışma kapsamında ele alınmıştır ve ara değerlendirme kullanarak bir-bit dönüşümü yoluyla piksel altı doğrulukta hareket kestirimi önerilmiştir. Yapılan deneyler ile önerilen yaklaşımın standart 1BT yöntemine göre başarımlarını artırdığı açıkça gösterilmiştir. Hatta 0.25 piksel doğruluk için 2BT yönteminden daha yüksek başarımlar elde edilmiştir. İleriki çalışmalarda orijinal piksel boyutundaki 1BT imgelerin, ikili halde ara değerlendirilmesi yaklaşımı değerlendirilebilir. Alt piksel seviyesinde 1BT'in tam arama işlem yükünü azaltmak için seyrek arama noktaları kullanan yaklaşımların yöntemle tümleştirilmesi denenebilir.

## 5. Kaynakça

- [1] B. Natarajan, V. Bhaskaran, and K. Konstantinides, "Low-Complexity Block-Based Motion Estimation via One-Bit Transforms," IEEE Trans. on Circuits and Systems for Video Technology, Vol. 7, No. 4, pp. 702-706, Aug. 1997.
- [2] P. H. W. Wong and O. C. Au, "Modified One-Bit Transform for Motion Estimation," IEEE Trans. Circuits Syst. Video Technol., Vol. 9, No. 7, pp. 1020-1024, Oct. 1999.
- [3] T. Koga, K. Iinuma, A. Hirano, Y. Iijima, and T. Ishiguro, "Motion-Compensated Interframe Coding for Video Conferencing," Proc. of Nat. Telecommun.Conf., pp. G5.3.1-5.3.5, Nov./Dec. 1981.
- [4] J. R. Jain and A. K. Jain, "Displacement Measurement and Its Application in Interframe Image Coding," IEEE Trans. Communication., Vol. COM-29, pp. 1799-1808, Dec. 1981.
- [5] R. Li, B. Zeng, and M. L. Liou, "A New Three-Step Search Algorithm for Block Motion Estimation," IEEE Trans. on Circuits and Systems for Video Technology, Vol. 4, No. 4, pp. 438-442, Aug. 1994.
- [6] A. M. Tekalp, Digital Video Processing. Englewood Cliffs, NJ: Prentice-Hall, 1995, pp. 432-455.
- [7] A. Ertürk and S. Ertürk, "Two Bit Transform for Binary Block Motion Estimation " IEEE Trans. On Circuits and Systems for Video Technology, Vol. 15, No. 7, pp. 938-946, July 2005.

- [8] G. Bjontegaard, "Motion Compensation with 1/4 pixel accuracy", ITU-T SG16/Q15, Geneva, Feb. 2000.



Şekil 3: 299 çerçeveden oluşan "Foreman" dizisinin ilk 80 çerçevesi ve  $8 \times 8$  piksel blok boyutu için, farklı hareket kestirimi yöntemleri ile elde edilmiş PSNR değerleri. (1BT= Bir-bit arama, 1BT (0.5)= Yarım piksel doğrulukta bir-bit arama), 1BT (0.25)= Çeyrek piksel doğrulukta bir-bit arama).

Tablo 1:  $16 \times 16$  piksel boyutunda bloklar için hareket kestirimi yöntemlerinin ortalama PSNR değerleri (dB)

YÖNTEM	VİDEO DİZİLERİ					
	Football (352x240) (125 çerçeve)	Flowergarden (352x240) (115 çerçeve)	Mobile (352x240) (140 çerçeve)	Tennis (352x240) (112 çerçeve)	Coastguard (352x288) (299 çerçeve)	Foreman (352x288) (299 çerçeve)
MAD	22.88	23.79	22.99	29.87	30.48	32.11
MAD(0.5piksel)	23.50	24.86	24.87	30.60	31.21	33.56
MAD(0.25piksel)	23.82	25.52	25.80	30.98	31.82	34.19
2BT	22.08	23.43	22.72	28.89	29.93	30.71
1BT	21.83	23.32	22.71	28.77	29.84	30.44
1BT (0.5piksel)	22.31	24.15	24.49	29.27	30.36	31.13
1BT (0.25piksel)	22.53	24.61	25.27	29.49	30.79	31.54

Tablo 2:  $8 \times 8$  piksel boyutunda bloklar için hareket kestirimi yöntemlerinin ortalama PSNR değerleri (dB)

YÖNTEM	VİDEO DİZİLERİ					
	Football (352x240) (125 çerçeve)	Flowergarden (352x240) (115 çerçeve)	Mobile (352x240) (140 çerçeve)	Tennis (352x240) (112 çerçeve)	Coastguard (352x288) (299 çerçeve)	Foreman (352x288) (299 çerçeve)
MAD	24.73	25.22	23.88	31.25	31.59	32.90
MAD (0.5 piksel)	25.78	26.88	26.30	32.23	32.61	34.72
MAD (0.25 piksel)	26.24	27.85	27.57	32.68	33.42	35.53
2BT	23.36	24.55	22.99	29.91	30.50	30.64
1BT	22.70	24.11	22.73	29.22	29.20	29.84
1BT (0.5 piksel)	23.38	25.09	24.73	29.71	29.64	30.87
1BT (0.25 piksel)	23.68	25.57	25.42	29.80	30.19	31.41