

# Videoda Nesne Sınıflandırması için Siluet Tabanlı Yöntem

Yiğithan Dedeoğlu<sup>1</sup>, B. Uğur Töreyn<sup>2</sup>, Uğur Güdükbay<sup>1</sup>, A. Enis Çetin<sup>2</sup>

Bilkent Üniversitesi  
Bilgisayar Mühendisliği Bölümü<sup>1</sup>  
Elektrik ve Elektronik Mühendisliği Bölümü<sup>2</sup>  
06800, Bilkent, Ankara

{yigithan, gudukbay}@cs.bilkent.edu.tr<sup>1</sup>,  
{bugur, cetin}@bilkent.edu.tr<sup>2</sup>

## Özetçe

Bu makalede akıllı gözetleme sistemleri oluşturulmasına yardımcı olacak, gerçek zamanda, videoda nesne sınıflandırması yapabilen bir örneğe dayalı makina öğrenme yöntemi sunulmaktadır. Sunulan yöntem sabit kamerayla izlenen bir alanda bulunan nesnelerin silüetlerinden yararlanarak nesnelere sınıflandırmaktadır. Nesne bölütlemesi için uyarlanabilir bir arka plan kestirim modeli kullanılmaktadır. Nesnelere önceden belirlenmiş insan, insan grubu, araç sınıflarına ayırmak için şablon eşleştirmeye dayalı bir güdümlü sınıflandırma yöntemi kullanılmıştır.

## 1. Giriş

Video kullanarak bir sahnede hareket eden nesnelere sınıflandırmak zor olduğu kadar birçok uygulama vaadeden doğurgan bir bilimsel problemdir. Bu problemi çalışmaktaki amacımız sıradan bir görüntüye dayalı gözetleme sistemine eklenebilecek gerçek zamanda nesne algılaması ve sınıflandırması yapabilecek bir sistem geliştirmektir. Zamana bağlı video bilgisinin çokluğunu ve karmaşıklığını göz önüne alırsak gerçek zamanda çalışacak bu sistemde kullanılacak algoritmaların ve metodların hızlı ve güvenilir çalışması gerekmektedir. Bu bildiride bu özelliklere sahip, sabit bir kamerayla elde edilen siyah beyaz video görüntüsü üzerinde çalışan bir sistem sunulmaktadır.

Sunulan sistemde, hareketli nesne tespiti iç ve dış mekanlarda başarıyla çalışan uyarlanabilir bir arka plan kestirimi metodu kullanılarak gerçekleştirilmektedir [6]. Hareketli pikseller sahnenin sabit arka planından bölütlendikten sonra birbirine bağlı bölgeler belirlenmekte ve bu bölgeler daha önceden tanımlanmış olan insan, insan grubu ve araç sınıflarına ayrılmaktadır. Nesne sınıflandırma algoritması tespit edilen hareketli nesnelerin silüetlerinin daha önceden sınıflandırılmış ve bir veritabanına kaydedilmiş şablonların silüetleriyle karşılaştırılmasına dayanmaktadır. Şablon veritabanı örnek videolardan örnek silüetlerin toplanması ve elle uygun sınıflara ayrılmasıyla oluşturulmuştur. Nesnelerin silüetleri birbirine bağlı bölgelerden bir çevrit işaretleme algoritması yardımıyla çıkarılmıştır [10].

Bildirinin geri kalan kısmı şu şekilde düzenlenmiştir. İlgili çalışmalar ikinci kısımda değerlendirilmiştir. Sonraki iki kısımda videoda hareketli nesne tespiti ve nesne sınıflandırması metodları anlatılmıştır. Deneysel sonuçlar beşinci kısımda ve sonuç altıncı kısımda sunulmuştur.

## 2. İlgili Çalışmalar

Nesne tespiti ve sınıflandırması konularında literatürde yapılmış birkaç inceleme makalesi yer almaktadır [1, 8, 22].

Videoda insanlara ve araçlara karşılık gelen hareketli bölgelerin bulunması sonraki analiz adımları için odaklanma sağlaması ve yapılacak işleri azaltması bakımından hemen hemen tüm görüntü işleme sistemlerinin ilk adımındır. Videoda hareketli nesne tespiti doğal sahnelerde meydana gelen ani ışık ve hava durumu değişimi ve karışıklığa neden olan tekrar eden hareketler (rüzgarda salınan ağaç yaprakları) gibi dinamik değişikliklerden dolayı güvenilir bir şekilde gerçekleştirilmesi zor olan bir problemdir. Hareketli nesne tespiti için sıklıkla kullanılan yöntemler arka plan kestirimi, istatistiksel metodlar, zamansal farklama ve optik akıştır [9, 11, 16, 20, 22].

Videoda tespit edilen hareketli bölgeler gerçek dünyada insanlara, araçlara, hayvanlara ya da karmakarışık hareket eden ağaç, çalı gibi farklı nesnelere karşılık gelir. Tespit edilen nesneyi düzgün bir şekilde takip etmek ve faaliyetlerini sağlıklı bir şekilde çözümlenmek için nesnenin tipini sınıflandırmak oldukça önemlidir. Halihazırda şekil tabanlı ve hareket tabanlı olmak üzere iki çeşit nesne sınıflandırma yöntemi vardır [22]. Şekil tabanlı metodlar nesnelerin kuşatan kutusu, alanı, silüetleri ve tespit edilen alanın gradyanı gibi iki boyutlu bilgilerden yararlanırken; harekete dayalı yöntemler nesnelerin zamansal olarak izlenen özelliklerinden faydalanırlar.

[14]'te anlatılan yöntemde tespit edilen nesnelerin silüetlerinin uzunluğu ve alanları kullanılarak nesnelere insan, araç ve diğer grubuna sınıflandırılmaktadır. Bu metod insanların genel olarak araçlardan daha küçük olmaları ve şekillerinin daha karmaşık olması varsayımına dayanmaktadır. Sınıflandırma ölçeği olarak dağılım kullanılmış ve nesnenin silüet uzunluğunun karesinin alanına oranı olarak tanımlanmıştır. Sınıflandırma her video çerçevesinde yapılmakta ve nesne takip algoritması sonuçları sınıflandırma sonuçlarının tutarlılığını iyileştirmede kullanılmaktadır.

Collins ve diğerleri tarafından geliştirilen sınıflandırma yöntemi tespit edilen nesnelerin bakışa bağlı görsel özelliklerini bir sinir ağı sınıflandırıcısını insan, insan grubu, araç ve diğer gruplarını tanımak üzere eğitmekte kullanılmıştır [6]. Sinir ağına giriş olarak dağılım, alan, nesnenin en-boy oranı ve kameranın zoom büyütme oranı kullanılmıştır. Önceki yöntemde olduğu gibi sınıflandırma her video çerçevesinde yapılmış ve sonuçların zamansal tutarlılığını sağlamak için bulunan sınıflar bir histogramda tutulmuşlardır.

## 3. Hareketli Nesne Tespiti

Önerilen yöntemde her video çerçevesinde bir ön plan haritası ve nesnelerin özelliklerini çıkarmak için arka plan modeli ve alt düzey görüntü işleme yöntemlerinin bileşimi kullanılmaktadır. Kullanılan arka plan kestirim modeli sabit bir kameradan gelen siyah-beyaz görüntü üzerinde çalışmaktadır ve bu modelin geliştirilmesinde [6]'da yapılan çalışmadan

esinlenilmiştir. Arka plan kestirimi modeli videonun ilk birkaç çerçevesini kullanarak bir referans arka planı oluşturur. Daha sonra yürürlükteki video çerçevesindeki her piksel değerinden, oluşturulan arka planda karşılık gelen piksel değeri çıkarılarak bir fark görüntüsü oluşturulur. Bu fark görüntüsü gürültü piksellerinin elenmesi için her piksel için uyarlanabilir bir eşik değeriyle filtrelenir. Referans arka plan resmi ve eşik değerleri dinamik değişikliklere uyarlanmak için bir sonsuz etki tepki (IIR) filtresiyle güncellenir.

$I_n(x)$ ,  $I$  video dizisinin,  $n$  zaman anındaki çerçevesinde,  $x$  konumundaki pikselin gri-düzeyi parlaklık değerini temsil etsin ve  $[0, 255]$  aralığında olsun.  $B_n(x)$  ise  $x$  konumundaki piksele karşılık gelen arka plan parlaklık değeri olsun ve  $I_0$ 'dan  $I_{n-1}$ 'e kadar olan video çerçevelerinden zamanla kestirilsin. Genel arka plan kestirimi yönteminin önerdiği üzere yürürlükteki video görüntüsünde  $x$  konumunda bulunan piksel aşağıdaki koşul sağlanıyorsa ön plana aittir:

$$|I_n(x) - B_n(x)| > T_n(x)$$

Burada  $T_n(x)$ ,  $I_0$ 'dan  $I_{n-1}$ 'e kadar olan video çerçevelerinden zamanla kestirilen uyarlanabilir eşik değeridir. Üstteki denklem 1'in ön planı, 0'ın arka planı temsil ettiği bir ikili ön plan haritası yaratmak için kullanılır. Bu ön plan haritasında değeri 1 olan pikseller hareketli alanları temsil eder. Referans arka plan görüntüsü olan  $B_n(x)$  ilk video görüntüsü olan  $I_0$  ile iklenendirilirken ( $B_0=I_0$ ), eşik değeri olan  $T_n(x)$ , daha önceden belirlenmiş olan bir sabit değerle (örn. 15) iklenendirilir.

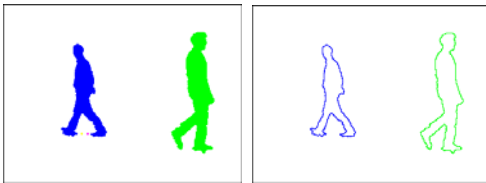
Bu sistem hem iç hem de dış mekanlarda kullanılacağı için, arka plan modelinin ani ışık değişimi ve uzun zamanlı arka plan değişikliği gibi dinamik değişikliklere kolay adapte olması gerekir. Bu amaçla referans arka plan görüntüsü ve eşik değerleri yeni gelen video çerçeveleriyle dinamik olarak güncellenir. Güncelleme yöntemi ön plana ait olarak belirlenen pikseller için ( $x \in FG$ ) ve arka plana ait olarak belirlenen pikseller ( $x \in BG$ ) için farklılık gösterir:

$$B_{n+1}(x) = \begin{cases} \alpha B_n(x) + (1 - \alpha)I_n(x), & x \in BG \\ \beta B_n(x) + (1 - \beta)I_n(x), & x \in FG \end{cases}$$

$$T_{n+1}(x) = \begin{cases} \alpha T_n(x) + (1 - \alpha)(\gamma \times |I_n(x) - B_n(x)|), & x \in BG \\ T_n(x), & x \in FG \end{cases}$$

Burada  $\alpha$ ,  $\beta$  ve  $\gamma$  ( $\in [0.0, 1.0]$ ) yeni gelen video görüntülerindeki bilginin ne kadarının arka plan ve eşik değeri görüntülerine ekleneceğini belirten öğrenme sabitleridir.

Üstteki algoritmanın sonucu bulunan ön plan görüntüsü genelde gürültü içerir ve bu yüzden özel artışlemlerden geçirilmeden sonraki basamaklarda kullanılmazlar. Morfolojik işlemlerden olan kemirme ve genişleme işlemleri gürültüyü temizlemek için ön plan piksel haritasına uygulanır ve gerçekte ön plana ait olmayan hatalı pikseller ile arka plana ait olmayan hatalı pikseller ayıklanır.



Şekil 1 Örnek nesnelere ve silüetleri

### 3.1 Nesne Özelliklerinin Hesaplanması

Ön plana ait bölgelerin belirlenmesinin ve gürültü piksellerinin temizlenmesi için artışlemlerin uygulanmasının ardından filtrelenen ön plan pikselleri [10]'da anlatılan iki aşamalı bağlı bileşen işaretleme algoritması kullanılarak bağlı bölgelere ayrılır ve etiketlenir. Nesnelere karşılık gelen ayrıntı bölgelerin bulunmasının ardından, kuşatan kutu, büyüklük, ağırlık merkezi ve silüet gibi nesne özellikleri hesaplanır.

$O$  nesnesinin ağırlık merkezini,  $C_m = (x_{C_m}, y_{C_m})$ , hesaplamak için aşağıdaki denklem kullanılır [17]:

$$x_{C_m} = \frac{\sum_i^n x_i}{n}, \quad y_{C_m} = \frac{\sum_i^n y_i}{n}$$

Burada  $n$ , nesneye karşılık gelen ön plan bölgesinde bulunan piksel sayısıdır.

Sınıflandırma algoritmasının hem eğitim hem test aşamasında nesnelere silüetleri ön plan piksel haritasından [10]'da anlatılan çevrit işaretleme algoritması yardımıyla çıkarılır. Şekil 1 tespit edilmiş örnek ön plan nesne bölgelerini ve bunlardan çıkarılan silüetleri gösterir. Nesnelere karşılık gelen bölgelerden çıkarılan bir diğer özellik de silüet uzaklık sinyalıdır.  $S = \{p_1, p_2, \dots, p_n\}$  bir  $O$  nesnesinin  $n$  noktadan oluşan ve ön plan piksel haritasında  $O$ 'ya karşılık gelen bölgenin en üst orta noktasından başlayarak elde edilmiş olan silüeti ve  $C_m$  de bu nesnenin ağırlık merkezi olsun. Bu durumda uzaklık sinyalı,  $DS = \{d_1, d_2, \dots, d_n\}$ , 1'den  $n$ 'e kadar olan her silüet noktasıyla,  $p_n$ , ağırlık merkezinin,  $C_m$ , arasındaki uzaklık,  $d_i$ , hesaplanarak şu şekilde oluşturulur:

$$d_i = Dist(c_m, p_i), \quad \forall i \in [1 \dots n]$$

Burada  $Dist$ , fonksiyonu iki a, b noktası arasındaki Euclid uzaklığıdır:

$$Dist(a, b) = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2}$$

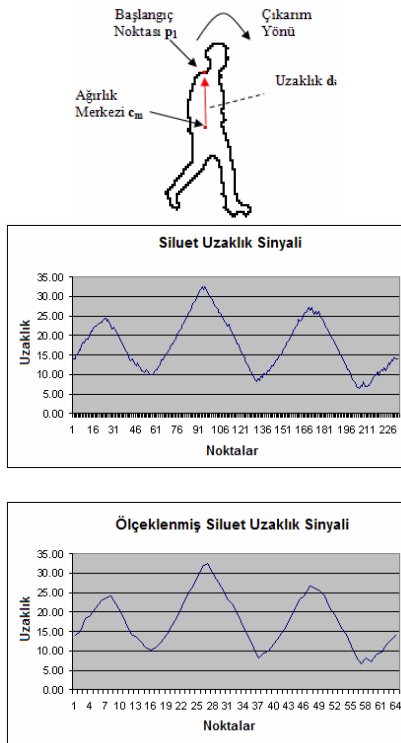
Videoda bulunan nesnelere farklı şekillere ve bu nedenle farklı uzunlukta silüetlere sahiptirler. Aynı nesne bile, çerçeveden çerçeveye farklılaşan çevrit uzunluğuna sahiptir. Farklı boydaki nesnelere karşılık gelen silüet uzaklık sinyallerini hatasız bir şekilde karşılaştırmak ve karşılaştırma ölçeğini büyüklük ölçeğinden değişimsiz hale getirmek için silüet uzaklık sinyalinin boyutu sabit tutulmaktadır.  $N$  uzaklık sinyalinin,  $DS$ , boyu ve  $C$ 'de değişmez sinyal uzunluğunu temsil eden bir sabit olsun. Bu durumda sabit uzunluktaki uzaklık sinyali olan  $\overline{DS}$ , alt ya da üst örnekleme yöntemiyle aşağıdaki şekilde hesaplanır:

$$\overline{DS}[i] = DS[i * \frac{N}{C}], \quad \forall i \in [1 \dots C]$$

Bir sonraki aşamada ölçeklenen uzaklık sinyali,  $\overline{DS}$ , birim alanı kapsaması için normalize edilir. Normalize edilen uzaklık sinyali aşağıdaki gibi hesaplanır:

$$\overline{DS}[i] = \frac{\overline{DS}[i]}{\sum_1^C \overline{DS}[i]}$$

Şekil 2 örnek bir silüeti ve orijinal ve ölçeklenmiş edilmiş bir silüet sinyallerini göstermektedir.



Şekil 2 Örnek silueti çıkarma, orjinal ve ölçeklenmiş siluet sinyalleri.

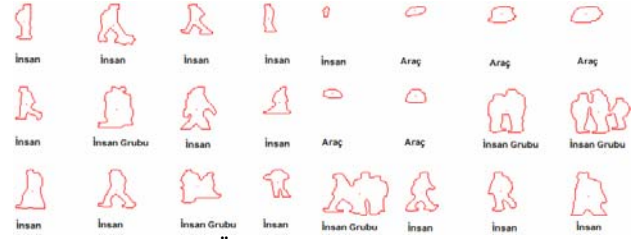
#### 4. Nesnelerin Sınıflandırılması

Birçok görüntüye dayalı gözetleme uygulamasının en büyük amacı daha üst seviye etkinlik analizi görevlerinde kullanılabilecek anlam bilgisinin videodan öğrenilmesidir. Tespit edilen nesnelerin tiplerinin belirlenmesi bu amaca ulaşmak için çok önemlidir. Nesne sınıfı bilgisinin yardımıyla daha özgül ve daha doğru sonuç veren üst seviye nesne etkinliği tespiti yapan yöntemler geliştirilebilir. Bu amaçla, görüntüye dayalı “akıllı” gözetim sisteminin bir parçası olarak kullanılabilecek, nesne şekil benzerliğine dayanan ve gerçek zamanda çalışan bir sınıflandırma metodu sunulmaktadır.

Tipik video sahneleri insan, araç, hayvan, doğal fenomen (örn. yağmur, kar) ve bitki gibi değişik nesnelere içerir. Ancak temelde gözetleme uygulamalarının ana hedefi insanlar ve araçlardır. Bu nedenle önerilen sistem nesnelere insan, insan grubu ve araç sınıflarına ayırmak için geliştirilmiştir.

Nesne sınıflandırma algoritması tespit edilen hareketli nesnelerin silüetlerinin daha önceden sınıflandırılmış ve bir veritabanına kaydedilmiş şablonların silüetleriyle karşılaştırılmasına dayanmaktadır ve iki adımdan oluşmaktadır.

- Çevrimdışı eğitim adımı: Nesne tipleri elle işaretlenerek bir örnek nesne silüet veritabanı oluşturulur.
- Çevrimiçi test adımı: Her video çerçevesinde bulunan nesnelerin silüetleri çıkarılır ve bu nesnelerin tipleri silüete dayalı özellikleri veritabanında önceden kaydedilmiş olan örnek şablon silüet özellikleriyle karşılaştırılarak gözetleme sırasında gerçek zamanda belirlenir. Karşılaştırma sonrasında veritabanında en az uzaklığa sahip bir şablon silüet bulunur. Bulunan bu silüetin tipi sınıflandırılmak istenen nesneye atanır.



Şekil 3 Örnek silüet şablon veritabanı

Şablon silüet veritabanı, farklı sahnelerden birçok nesnenin silüeti çıkarılarak oluşturulur. Sınıflandırma yöntemi nesne şekil benzerliğinden faydalandığı için veritabanında bulunan örnekler nesnelerin değişik pozlarını temsil edecek şekilde olmalıdır. Şekil 3, değişik pozlarda insan, insan grubu ve araç silüetlerinden oluşan 24 şablonluk bir örnek veritabanı göstermektedir.

Sınıflandırma adımında bu yöntem ham silüet bilgisini değil, çevrilmiş olan normalize edilmiş uzaklık sinyalini kullanmaktadır. Bu sayede veritabanında sadece normalize edilmiş uzaklık sinyali ve ilgili sınıf bilgisi saklanmaktadır, ve hem kayıt alanı açısından hem de işlemsel hız açısından verim sağlanmaktadır.

#### 4.1 Sınıflandırma Ölçüvi

Önerilen sınıflandırma ölçüvi nesne şekillerinin benzerliğine dayanmaktadır. Literatürde şekilleri karşılaştırmak için kullanılan sayısız yöntem bulunmaktadır [18, 4, 17, 2, 12] ve bunların dışında [21, 15] de farklı tekniklerin tartışılması açısından önemlidir.

Sunulan sınıflandırma ölçüvi A ve B nesnelere arasındaki benzerliği, bu nesnelerin uzaklık sinyalleri ( $DS_A$  ve  $DS_B$ ) arasındaki uzaklığı hesaplayarak bulmaktadır. İki ölçeklenmiş ve normalize edilmiş sinyal ( $DS_A$  ve  $DS_B$ ) arasındaki uzaklık aşağıdaki şekilde hesaplanmaktadır:

$$Dist_{AB} = \sum_{i=1}^n |DS_A[i] - DS_B[i]|$$

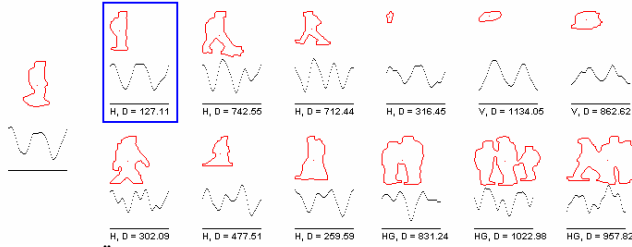
Bir O nesnesinin sınıfını,  $T_O$ , bulmak için bu nesnenin uzaklık sinyali ( $DS_O$ ) ile şablon veritabanındaki (TDB) tüm uzaklık sinyalleri karşılaştırılır. P şablon nesnesinin tipi olan  $T_P$ , sınıfı aranan O nesnesinin tipi olarak atanır,  $T_O = T_P$ . Bunun için P nesnesinin aşağıdaki koşulu sağlaması gerekir:

$$Dist_{OP} \leq Dist_{OB}, \forall P \in TDB$$

Şekil 4'te örnek bir sorgu nesnesi ve şablon veritabanı nesnelerinin silüetleri ve uzaklık sinyalleri gösterilmektedir. Sınıflandırmanın sonucunu iyileştirmek ve tutarlılığını artırmak için atanan nesne tipleri bir histogramda tutulmakta ve her çerçevede her nesnenin ilgili histogramındaki en çok sayıya sahip olan sınıf nesnenin sınıfı olarak seçilmektedir.

#### 5. Deneysel Sonuçlar

Tüm testler görüntü işleme algoritmalarını geliştirmek için kullandığımız bir video oynatma uygulamasında ve Microsoft Windows XP Professional işletim sistemi çalıştıran Intel PIV-2600 MHz işlemcili ve 512MB bellekli bir bilgisayarda gerçekleştirilmiştir.



**Şekil 4 Örnek sorgu objesi ve şablon veritabanındaki birçok nesneye olan uzaklığı (D). Nesne tipleri İnsan (H), İnsan Grubu (HG) ve Araç (V). Eşleşen obje kuşatan kutuyla gösterilmiştir.**

Nesne sınıflandırma algoritmasını test etmek için öncelikle silüetleri çıkarmak ve elle sınıflandırmak için geliştirdiğimiz bir uygulama kullanılarak bir örnek şablon nesne veritabanı oluşturuldu. Bu eğitime işleminde içinde hareketli insan, insan grubu ve araç görüntüleri bulunan dört video dizisi kullanıldı. Oluşturulan şablon nesne veritabanı kullanılarak yine içinde insan, insan grubu ve araçlar bulunan video dizileri test için kullanıldı. Test sonuçlarıyla oluşturulan hata matrisi Tablo 1 de gösterilmektedir.

	İnsan	İnsan Grubu	Araç	Başarı Yüzdesi
İnsan	175	13	20	84.13%
İnsan Grubu	12	52	14	66.67%
Araç	38	22	238	79.86%
Ortalama Başarı				76.88%

**Tablo 1 Nesne sınıflandırması testleri sonucu oluşturulan hata matrisi.**

## 6. Sonuç

Bu çalışmada videoda gerçek zamanda hareketli nesne takibi ve sınıflandırması yapabilen bir sistem geliştirilmiştir. Test sonuçları hatalı kararlardaki yüzdenin sınıflandırma algoritmasında yapılacak geliştirmelerle aşağı çekilmesiyle sistemin gerçek zamanda çalışacak gözetleme sistemlerinde kullanılabileceğini göstermektedir. Sunulan sistemin en büyük avantajı gerçek zamanda çalışabilmesidir. Ayrıca bu sistem multimedia veritabanlarında otomatik veri oluşturulması amacıyla çevrim dışı olarak da kullanılabilir.

Yöntemin kameranın bakış açısına dayalı olması oluşturulan şablon nesne veritabanının her bakış açısında kullanılamamasına neden olmaktadır. Bu problem eğitim verisinin otomatik olarak oluşturulmasıyla (k-ortalama yöntemiyle gruplandırma gibi) giderilebilir.

## Kaynakça

- [1] J.K. Aggarwal and Q. Cai. Human motion analysis: a review. *Computer Vision and Image Understanding*, 73(3):428–440, March 1999.
- [2] E.M. Arkin, L.P. Chew, D.P. Huttenlocher, K. Kedem, and J.S.B. Mitchell. An efficiently computable metric for comparing polygonal shapes. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 13:209–216, 1991.
- [3] O. Chomat, J.L. Crowley, Recognizing motion using local appearance, *International Symposium on Intelligent Robotic Systems*, University of Edinburgh, 1998.
- [4] R. T. Collins, R. Gross, and J. Shi. Silhouette-based human identification from body shape and gait. In *Proc. of Fifth IEEE Conf. on Automatic Face and Gesture Recognition*, pages 366–371, 2002.

- [5] R. Cutler and L.S. Davis. Robust real-time periodic motion detection, analysis and applications. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 8, pages 781–796, 2000.
- [6] R. T. Collins et al. A system for video surveillance and monitoring: VSAM final report. Technical report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, May 2000.
- [7] Y. Dedeoglu, Moving object detection, tracking and classification for smart video surveillance, Master's Thesis, Dept. of Computer Eng. Bilkent University, Ankara, 2004.
- [8] D. M. Gavrilu. The analysis of human motion and its application for visual surveillance. In *Proc. of the 2nd IEEE International Workshop on Visual Surveillance*, pages 3–5, Fort Collins, U.S.A., 1999.
- [9] I. Haritaoglu, D. Harwood, and L.S. Davis. W4: A real time system for detecting and tracking people. In *Computer Vision and Pattern Recognition*, pages 962–967, 1998.
- [10] F. Heijden. Image based measurement systems: object recognition and parameter estimation. Wiley, January 1996.
- [11] J. Heikkilä and O. Silven. A real-time system for monitoring of cyclists and pedestrians. In *Proc. of Second IEEE Workshop on Visual Surveillance*, pages 74–81, Fort Collins, Colorado, June 1999.
- [12] H.Ramoser, T.Schlgl, M.Winter, and H.Bischof. Shape-based detection of humans for video surveillance. In *Proc. of IEEE Int. Conf. on Image Processing*, Barcelona, Spain, 2003.
- [13] A. J. Lipton. Local application of optic flow to analyse rigid versus non-rigid motion. Technical Report CMU-RI-TR-99-13, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, December 1999.
- [14] A. J. Lipton, H. Fujiyoshi, and R.S. Patil. Moving target classification and tracking from real-time video. In *Proc. of Workshop Applications of Computer Vision*, pages 129–136, 1998.
- [15] S. Loncaric. A survey of shape analysis techniques. *Pattern Recognition*, 31(8):983–1001, August 1998.
- [16] A. M. McIvor. Background subtraction techniques. In *Proc. of Image and Vision Computing*, Auckland, New Zealand, 2000.
- [17] E. Saykol, U. Gudukbay, and O. Ulusoy. A histogram-based approach for object-based query-by-shape-and-color in multimedia databases, *Image and Vision Computing*, vol. 23, No. 13, pages 1170–1180, November 2005.
- [18] E. Saykol, G. Gulesir, U. Gudukbay, and O. Ulusoy. KiMPA: A kinematics-based method for polygon approximation. In *Lecture Notes in Computer Science (LNCS)*, Vol. 2457, pages 186–194, *Advances in Information Sciences (ADVIS'2002)* Edited by Tatyana Yakhno, Springer-Verlag, 2002.
- [19] C. Schuldt, I. Laptev and B. Caputo, Recognizing human actions: a local SVM approach, In *Proc. of ICPR'04*, Cambridge, UK.
- [20] C. Stauffer and W. Grimson. Adaptive background mixture models for realtime tracking. In *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, page 246252, 1999.
- [21] R.C. Veltkamp and M. Hagedoorn. State-of-the-art in shape matching, pages 87–119. *Principles of Visual Information Retrieval*. Springer, 2001.
- [22] L. Wang, W. Hu, and T. Tan. Recent developments in human motion analysis. *Pattern Recognition*, 36(3):585–601, March 2003.
- [23] L. Wixson and A. Selinger. Classifying moving objects as rigid or non-rigid. In *Proc. of DARPA Image Understanding Workshop*, pages 341–358, 1998.