

## 1.Giriş

Kimlik belirlemede ve erişim kontrolü için biyometriklerin kullanıldığı günümüzde yürüyüş (gait) analizinin bu alanda aktif bir araştırma sahası yaratması biraz sürpriz oluyor. (Sürpriz mi oluyor yoksa sürpriz olmuyor mu?) Bilgisayara bakış açısıyla düşünüldüğünde insan yürüyüşünün kişi tanımda uygun bir işaret olmasını sağlayan iki önemli neden var. İlki, parmak izi ve retina taramasının aksine, yürüyüşün uzak bir mesafeden analiz edilebilmesidir. İkincisi ise, hiçbir iki vücut hareketinin aynı olmasa da, yürüyüşün her birey için karakteristik bir biyometrik olmasıdır. Bu alanda ki bilimsel merak Fransız fizikçi ve fotoğrafçı olan Etienne-Jules Marey'e(1830-1904) uzanmaktadır. İlk defa 1920'lerde Moskova'da Nikolai A. Bernstein tarafından yapılan klinik çalışmalarda yürüyüşün her birey için karakteristik (eşsiz,kişiye has) olduğunu açığa çıkardı. Bu bulgular batıda bilinmesede, bu bulgular 1960'larda tıp çalışmalarıyla desteklenip kuvvetlendirildi[murray67]. Daha sonra 1970'lerde yapılan psikolojik deneylerde, insanların kolayca kişileri yürüyüşlerinden tanıyabildiği ortaya konuldu.[cutting77]. Tıpta kullanım geçmişi düşünüldüğünde, bilgisayar ortamında yürüyüş analizlerin yapılması henüz başlangıç aşamasındadır denilebilir. Buna rağmen, görsel tabanlı yürüyüş analizi için ortaya çıkarılmış yöntemler çok geniş bir yelpazeyi kapsıyor. Yakın geçmişte yapılmış, ayrıntılandırılmış farklı yaklaşımlar [moeslund00,sarkar05]'de incelenmektedir. Bu araştırmalardan görüldüğü üzere, bazı yaklaşımlar insan silüetini resim kesitlerinden çıkartmasa da [niyogi94,cutler00], yeni yapılan çalışmaların büyük bir çoğunluğu şekil analizi üzerine yoğunlaşmıştır. Little ve Boyd[little98] hareketli nesneleri **optical flow**'dan bulmakta ve karakteri özellikleri(feature)

## 1. Introduction

In times of increased interest in biometrics for authentication and access control, it comes as little surprise that gait analysis has become an active area of research. From a computer vision point of view, there are two major reasons why human gait is an appealing cue for person identification: in contrast to fingerprint- or retina scans, gait can be analysed from a distance. Yet, although no two body movements are ever the same, gait is as characteristic of an individual as other biometrics. Scientific interest in this curiously familiar fact dates back to the French physician and photographer Etienne-Jules Marey (1830–1904). In the 1920s, Nikolai A. Bernstein conducted first clinical studies in Moscow, which revealed that gait is a unique personal trait. Unknown in the West, these findings were corroborated by medical studies from the 1960s [murray67]. Subsequent psychological experiments conducted in the 1970s revealed that humans effortlessly recognise people by the way they walk [cutting77].

Compared to its tradition in medicine, the history of human gait as a research topic in computer vision is rather brief. Nevertheless, the spectrum of methods that have been proposed for vision-based gait analysis is already vast. Recent and detailed surveys of different approaches can be found in [moeslund00,sarkar05]. From these surveys it becomes apparent that, although some contributions to the problem did not extract human silhouettes from image sequences [niyogi94,cutler00], the majority of recent work relies on shape analysis. Little and Boyd [little98] derive the shape of motion from optical flow and use this to compute characteristic features.



Fig. 1. Silhouettes extracted from videos of walking people. Apparently, the information contained in sequences of shapes allows to distinguish between *normal* and *abnormal* gait. (a) Person slowly walking away from the camera; (b) Person wavering towards the camera.

çıkartmaktadır. Bobick ve Johnson[bobick01], yürüyen insan eğer (binary) resmini alt bölümlere bunu kullanarak ayırıp, uzuvları arasında ki farklı ilişkileri ölçmüştür. Benzer biçimde, Lee ve Grimson[lee02], silüetleri elipslerle benzeterek, onlardan farklı özellikler çıkartmıştır. Veres ve arkadaşları[veres04], varyans teknikleri uygulayarak, ikili değer resminde hangi bölgelerin yürüyüş hakkında en karakteristik bilgileri verdiğini bulmuştur. Collins ve arkadaşları[collins02], **propose template matching of body silhouettes as a baseline algorithm in gait recognition.??**. Sarkar ve arkadaşları [sarkar05] geçici şekil istatistiklerini(**temporal shape statistics?**) kullanmıştır. Tolliver ve Collins[tolliver03], **store normalised silhouettes as vertices of similarity graphs and recover significant parts of the walking cycle from computing the eigenvectors of the Laplacian matrix of these graphs. BenAbdelkader et al. [abdelkader04] compute self-similarity plots from silhouettes and apply subspace methods to recognise individuals.**

Diğer makale yazarları yaptıkları işleri Kendal[kendal84], Boyd[boyd04], Veeraraghavan ve arkadaşları[Veeraraghavan04] ve Wang ve arkadaşlarının geliştirdiği istatistiksel şekil teorisi oturtmuştur. Bu çalışmalarda insan silüetleri **Procrustean uzaklığı(Nedir?)** uygulanarak karşılaştırılır ve sınıflandırılır.

Rağbette olan şekil tabanlı yaklaşımların dışında, yukarıda özeti geçilen bu makaleler, yürüyüş analizlerinin hemen hepsinin kimlik saptamada(**identification, tanımlama da denebilirdi**) kullanıldığını ortaya koymaktadır. Halbuki, yürüyüş cabundan daha fazlasını açığa çıkartmaktadır. Yapılan psikolojik deneylerde, insanların kişilerin silüetlerine ve ya hareketli ışık gösterimlerine (eklemlerine ışık yerleştirilmiş kişilerin videolarından) bakarak kişilerin hareket

Bobick and Johnson [bobick01] subdivide the binary silhouette of walking people into body parts and measure different relations among limbs. Similarly, Lee and Grimson [lee02] approximate

silhouettes by ellipses and compute feature descriptors therefrom. Veres et al. [veres04] apply analysis of variance techniques to determine which parts of a binary shapes provide the most characteristic information for gait analysis. Collins et al. [collins02] propose template matching of body silhouettes as a baseline algorithm in gait recognition. Sarkar et al. [sarkar05] favour temporal shape statistics. Tolliver and Collins [tolliver03] store normalised silhouettes as vertices of similarity graphs and recover significant parts of the walking cycle from computing the eigenvectors of the Laplacian matrix of these graphs. BenAbdelkader et al. [abdelkader04] compute self-similarity plots from silhouettes and apply subspace methods to recognise individuals. Other authors base their work on a statistical theory of shape developed by Kendal [kendal84], Boyd [boyd04], Veeraraghavan et al. [Veeraraghavan04] and Wang et al. [wang02] apply Procrustean distances to compare and classify human silhouettes.

Apart from the popularity of shape-based approaches, the papers considered in this brief survey reveal that gait analysis is almost exclusively being applied to identification tasks. However, gait can disclose more than identity. In psychological experiments, human subjects easily recognise different types of motion (e.g., walking, dancing, wavering, etc.) when asked to interpret moving light displays (obtained from filming subjects wearing light bulbs on their joints) or sequences of body silhouettes [johansson84].

dizisine (series) bakarak doğrulayabilir. Bu resimdeki şekiller yürüyen insanların videolarından çıkartılmıştır. Bir kişinin yavaşça kameradan uzaklaştığını gösteren Şekil-1(a)'da ki dizide olağan dışı bir şey yok iken, Şekil-1(b)'de gösterilen dizideki kişinin yürümekte zorlandığı kolayca anlaşılmaktadır. Bu diziden yakalan hareketler kişinin şüpheli (dubious) bir durumda olduğunu ileri sürmektedir.

tiplerini ( yürüme, dans etme, yalpalama..vb) kolaylıkla algılayabildikleri görülmüştür[johansson84]. Bu iddia'yı okuyucular Şekil-1'deki iki farklı yürüyüş Olağan dışı yürüyüşlerin tespiti için yazılmış sayısız araştırma olsa da, bu teknikleri incelemek zaman kaybı olmayacaktır. Aşağıda sunulan çalışma bu yönde bir ilk adımı temsil etmektedir. İlk olarak, hızlı ve doğru sonuçlar veren, yeni geliştirilmiş (novel, orijinal de olabilir?) bir şekil kodlama (shape encoding) tasarımı gösterilmiştir. Bu tasarım basit olarak şekillerin bir vektör uzayına gömülmesini sağlamaktadır. Bu sayede vektör olarak tanımlanan verilerin sınıflandırılması sağlayan birçok örüntü tanıma ve machine learning(?) tekniklerinden yararlanma olanağı doğar. 4. Bölüm'de, yürüyüşün destek vektör sınıflandırılmasıyla elde edilmiş sonuçlarını göstermekteyiz. which demonstrates that our approach allows for characterising types of movements across individuals. Bir özet ve genel görüş, makaleyi tamamlamaktadır.

## 2. Güçlü bir şekil kodlama

Videodan olağan dışı yürüyüş tespiti için yaptığımız yaklaşım, aşağıda bahsettiğimiz düşünceler üzerine ilerlemiştir. Şekil-1'den anlaşılacağı üzere, şekil dinamikleri, olağan ve olağan dışı yürüyüşlerin birbirinden ayrılmasında önemli bir işaret olmaktadır. Bu yüzden, giriş bölümünde açıklanan genel eğilimi takip edip, yürüyüş sınıflandırmayı şekil tabanlı yöntemler kullanarak yapmak uygun olacaktır. En basit olarak düşünüldüğünde, olağan dışı yürüyüş tespiti iki sınıflı bir problem olarak ele alınabilir. Bu yüzden, problemin çözümünde destek vektör makinası(DVM) kullanılabılır. DVM'lerin iki sınıflı sınıflandırmalarda güvenilir olarak çalıştığı bilinmektedir. Ancak DVM'lerinin kullanılması, analiz yapılacak silüetlerin vektör uzayına gömülmesini gerektirmektedir.

The reader may verify this claim by considering the two series of silhouettes shown in Fig. 1. The shapes in this figure were extracted from videos of walking people.

While there is nothing unusual in the series in Fig. 1(a), which shows a subject slowly moving away from the camera, the subject in the series in Fig. 1(b) obviously has difficulties walking. The movements captured in this sequence suggest the person is in a dubious condition.

As there are numerous applications for the detection of abnormal gait, it seems worthwhile to explore techniques that can accomplish this. The work presented in the following is a first step in this direction. First, we describe a novel, fast and robust shape encoding scheme. We show that this encoding scheme provides a straightforward vector space embedding of shapes. Therefore, it leads to a rich choice of pattern recognition and machine learning techniques that deal with the classification of vectorial data. In Section 4, we present experimental results obtained from support vector classification of gait, which demonstrates that our approach allows for characterising types of movements across individuals. A summary and an outlook completes this contribution.

## 2. Robust shape encoding

Our approach to detecting abnormal gait from video was guided by the following considerations. Fig. 1 indicates that shape dynamics provide a strong cue for distinguishing normal from abnormal gait. Therefore, it appears natural to follow the general trend identified in the introduction and to base gait classification on shape analysis. In its most rudimentary form, the classification task of detecting abnormal gait can be treated as a two-class problem. Therefore, it might be solved using a support vector machine approach, because support vector machines are known to reliably accomplish two-class classification. Applying support vector machines, however, requires a vector space embedding of the silhouettes we are concerned with. In contrast to the identification task, abnormal gait detection should abstract from personal traits.

Kimlik tanıma işleminin aksine, olağan dışı yürüyüş tespiti kişisel özelliklerden soyutlanmıştır. (In contrast to the identification task, abnormal gait detection should abstract from personal traits.) İhtiyaç duyulan vektör uzayı gömülümü bu yüzden vücut silüetinin genel durumunu yansıtmalıdır.

Ideally, it also would be insensitive to noisy boundaries and would not require much computational effort. Aşağıdaki şekil kodlama tasarısı verilen bu kriterlerin hepsine uymaktadır.

Bazı segmentasyon işlemlerinde elde edilen Şekil-1'deki ikili resimlerden (binary image), bir  $S$  şeklinin  $L$  adet pikselden oluşan bir dizi olduğunu anlıyoruz  $\mathcal{S} = \{\vec{p}_k \in \mathbb{R}^2 | k = 1, \dots, L\}$ . Şekil-2 aşağıda anlatılan işlemi görsel olarak ortaya koymaktadır. Bu işlem  $m \times n$  'lik  $B_{ij}$  kutularından oluşan  $\mathbf{B}$  dizisini (array) hesaplamaktadır. Bu, şeklin kaba bir temsili olarak düşünülebilir.

a)  $S$  piksel dizisinden,  $B(S)$  sınırlandırılmış kutuları (bounding box) hesapla,

b) Bunları  $n$  adet eşit genişlikteki dikey dilimlere ayır,

c) Sınırlandırılmış kutular olan  $B(S_j)$ 'leri oluşturulmuş her  $S_j$  piksel seti için hesapla ( $j=1, \dots, n$ )

d) Her  $B(S_j)$ 'yi  $m$  adet yatay eşik uzunluktaki dilimlere ayır;

e) Sınırlandırılmış kutular olan  $B_{ij} = B(S_{ij})$ 'leri her bir oluşturulmuş piksel seti için hesapla. ( $i=1, \dots, m$ );

Açıkça görülmektedir ki bu işlem  $mn$ 'de doğrusaldır (?). İnterpiksel ilişkilerin hesaplanmasına ve tekrar edilen maksimizasyon (minimizasyon) basamaklarına gerek yoktur (?). Bu yüzden, bu işlemin ortalama işlem karmaşıklığı  $O(mn\hat{p})$  'dir.  $\hat{p}$ ,  $B(S_{ij})$  kutusunun ortalama piksel sayısını vermektedir.

The required vector space embedding should therefore capture general aspects of body silhouettes. Ideally, it also would be insensitive to noisy boundaries and would not require much computational effort. The shape encoding scheme presented in this section meets all these criteria. Given binary images, as shown in Fig. 1, produced by some segmentation process, we understand a shape  $S$  to be a set of  $L$  pixels,  $\mathcal{S} = \{\vec{p}_k \in \mathbb{R}^2 | k = 1, \dots, L\}$ . Fig. 2 visualises the following procedure that computes an  $m \times n$  array  $B$  of boxes  $B_{ij}$ , which can be thought of as a coarser representation of a shape:

(a) compute the bounding box  $B_{SS}$  of a pixel set  $S$

(b) subdivide it into  $n$  vertical slices of equal width;

(c) compute the bounding boxes  $B(S_j)$  for each of the resulting pixel sets  $S_j$  where  $j = 1, \dots, n$ ;

(d) subdivide each  $B(S_j)$  into  $m$  horizontal slices of equal height;

(e) compute the bounding boxes  $B_{ij} = B(S_{ij})$  for each of the resulting pixel sets  $S_{ij}$  where  $i = 1, \dots, m$ ;

Obviously, this procedure is linear in  $mn$ . It does not require the computation of interpixel relations, and there is no need for repeated maximisation (or minimisation) steps. Therefore, the procedure's average computational complexity is  $O(mn\hat{p})$ , where  $\hat{p}$  denotes the average number of pixels in a box  $B(S_{ij})$ .

Ayrıca, Şekil-2(e)'deki her bir  $B(S_{ij})$  kutusu,  $(x_{ij}, y_{ij})$  koordinatlarındaki  $h_{ij}$  uzunluğu ve  $W_{ij}$  genişliğindeki genelleştirilmiş bir piksel olarak gösterilebilir (Şekil 3). Bu kaba olarak yaratılmış yeni şeklin hafızada kaplayacağı yer  $4mn$  olmaktadır (mxn adet piksel, her pikselin x-y koordinatları boy ve yüksekliğini temsil eden 4).  $m$  ve  $n$ 'nin küçük değerleri için, bu tekrar eden (iterative) kutulara bölme işlemi, hızlı ve hafıza tasarrufu sağlayan bir şekil çıkarımı (abstraction) oluşturmaktadır. Şekil-3'e bakıldığında makul büyüklükteki bir  $m \times n$  dizi boyutu yeteri kadar hassas (accurate) bir temsil sağlamaktadır. Bu göreceli açıklama Şekil-5'teki grafiklerle nesnelleştirilmektedir. 6 kişinin yürüyüş videosundan çıkarılmış olan, ortalama 18878 pikselden oluşan 1178 silüetle yapılan hesaplamaların özetini göstermektedir. Verilen  $h$  yüksekliği ve  $w$  genişliğindeki ilk sınırlandırılmış kutudan (initial bounding box) dizi boyutları  $m$ ,  $n$ 'nin bir fonksiyonu olarak hesaplanmıştır (bkz: Şekil-4):

$$m(n) = \left\lfloor \frac{h}{w} n \right\rfloor \quad (1)$$

$\lfloor x \rfloor$ ,  $x$ 'in en yakın küçük doğal sayıya yuvarlanmış halidir, örneğin,  $\lfloor x \rfloor = \sup \{ y \in \mathbb{N} \text{ ve } y \leq x \}$ .

Şekli sınırlandırılmış kutulara ayırma işleminden sonra elde edilmiş olan ortalama sıkıştırma oranı  $1-2mn/L$  Şekil-5(a)'da gösterilirken, Şekil-5(b)'de,  $S$  şekli ve onun

Moreover, each box  $B(S_{ij})$  as in Fig. 2(e) can be treated as a generalised pixel of height  $h_{ij}$  and width  $W_{ij}$  at location  $(x_{ij}, y_{ij})$  as in Fig. 3. The storage requirement of a coarse shape representation is therefore a mere  $4mn$ . For small values of  $m$  and  $n$ , this iterative process of bounding box splitting yields a fast and storage efficient abstraction of shapes. In addition, Fig. 3 also exemplifies that already moderate array dimensions  $m \cdot n$  can produce fairly accurate representations. This subjective impression is objectified by the graphs in Fig. 5. They summarise measurements obtained from 1178 silhouettes of an average size of 18878 pixels that were extracted from 6 video sequences of walking people. Given the height  $h$  and width  $w$  of a shape's initial bounding box, the array dimension  $m$  was computed as a function of  $n$  (see Fig. 4):

$$m(n) = \left\lfloor \frac{h}{w} n \right\rfloor \quad (1)$$

where  $\lfloor x \rfloor$  indicates rounding  $x_R$  to the nearest lower integer, i.e.,  $\lfloor x \rfloor = \sup \{ y \in \mathbb{N} \text{ ve } y \leq x \}$ .

While Fig. 5(a) shows the average compression rate  $1-2mn/L$  obtained in shape encoding by bounding box splitting, Fig. 5(b) depicts the normalised Hamming distance  $D = dH(S, B)/L$  between a shape  $S$  and its coarse representation as a box array  $B$ .

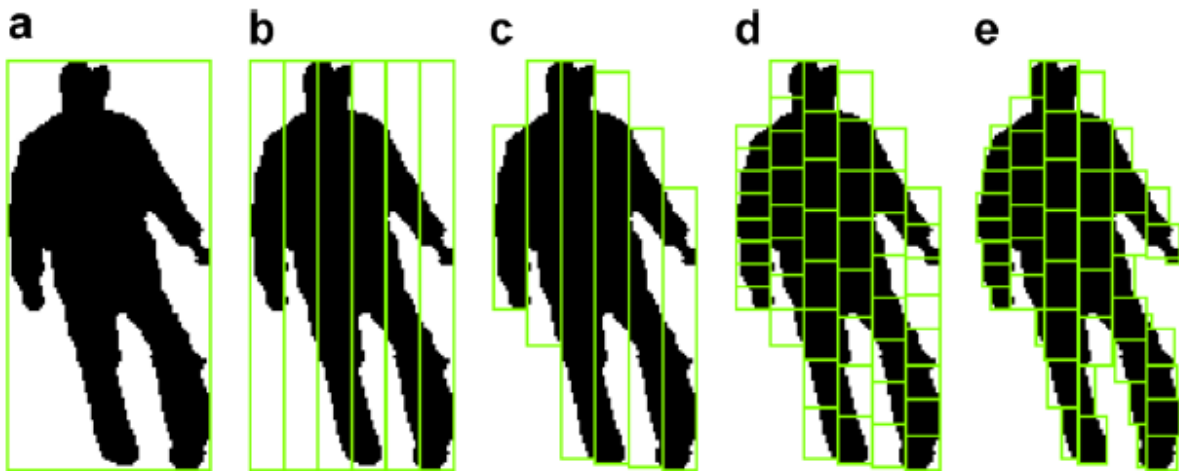


Fig. 2. Example of using bounding box splitting to map a  $8 \times 6$  lattice onto a shape.



kaba temsili olan **B** kutu dizisinin normalize edilmiş Hamming uzaklığı ( $D = dH(S, \mathbf{B})/L$ ) tasvir edilmiştir. Dikkate alınması gereken bir nokta; sıkıştırma oranı, büyüyen  $n$  değerleri için yavaşça azalırken, normalize edilmiş yeniden yapım hatası (**reconstruction error**) hızlıca azalmaktadır. %93'lük bir sıkıştırma oranında çoktan %4'ün altına inmiştir. Dahası, Şekil-5(b)'deki hata çizgilerinden, aynı sıkıştırma oranında, hatanın varyansının neredeyse kaybolduğunu görebiliriz. Bu yüzden, küçük boyutlardaki kutu dizilerinde bile birçok durum (**case**) için yeterince bilgi alabileceğimiz silüet temsilleri elde edilmektedir. Bu temsiller bize birçok durumda yüksek hassasiyette sonuçlar aldırılmaktadır. Sınırlandırılmış kutulara ayırma işleminin çalışma süresi korkulacak değerlerin çok uzağında kalmaktadır ve gerçek zamanlı video işlemleri yapmaya uygundur [bauckhage05].

Buna ilaveten, bu yöntem basit bir geometrik hesaplamayı esas alsa da, 2 boyutlu(2D) şekiller ve 2D'lu kafesler arasında bir benzer şekilliliği (**homomorphism**) göstermektedir (?). Eğer dizinin sol alt kısmından başlayarak bağımsız kutulardan oluşan kafes koordinatlarını ( $i, j$ ) saymaya başlarsak, o zaman bu kutulardan ( $B(S_{ij})$ ), bir diğerinin altında olanlarının kafes koordinatlarından  $i$ , daima daha küçük bir kafes koordinatı olur. Ayrıca, bir diğerinin sol tarafında olan kutuların kafes koordinatlarından  $j$ , daima küçük olur (bknz: Şekil-6) (ANLATILMAK İSTENEN NE ?). Bu topoloji çerçevesinde, kutu dizisinde temsil edilen şekillerin doğasını koruyarak kolayca bu şekilleri vektör uzayına gömebiliriz.

Eğer  $v$  ilk(initial) sınırlandırılmış kutu  $S$ 'nin (bölünmeden önceki mi ?) sol alt köşesinin bulunduğu noktayı,  $w$  ve  $h$  sırasıyla genişliğini ve uzunluğunu ve  $u_{ij}$ ,  $B(S_{ij})$  kutusunun orta noktasını belirttiği takdirde;

$$\vec{\mu}^{ij} = \begin{pmatrix} (u_x^{ij} - v_x)/w \\ (u_y^{ij} - v_y)/h \end{pmatrix} \quad (2)$$

Note that while the compression rate decreases slowly for growing values of  $n$ , the normalised reconstruction error decreases quickly. At a compression rate of about 93%, it already drops below 4%. Moreover, from the error bars in Fig. 5(b), we see that at the same level of compression the variance of the error already almost vanishes. Therefore, even box arrays of small sizes provide descriptions that capture the essential properties of a silhouette and are accurate across a wide variety of cases. The runtime required for bounding box splitting will be far from threatening real time constraints that are important in most video-based applications [bauckhage05].

Furthermore, note that this simple method realizes a homeomorphism between 2D shapes and 2D lattices even though it only relies on basic computational geometry. If we begin counting the lattice coordinates ( $i, j$ ) of individual boxes starting with the lower left of an array, then boxes  $B(S_{ij})$  situated below other boxes will always have smaller lattice coordinates  $i$ . Also, lattice coordinates  $j$  of boxes to the left of other boxes are always smaller (see Fig. 6). Due to this topology preserving nature of the box array representation of shapes, it is straightforward to embed shapes in a vector space. If the vector  $\sim v$  denotes the location of the bottom left corner of the initial bounding box of  $S$ ,  $w$  and  $h$  denote its width and height and the vector  $\sim u_{ij}$  denotes the centre of box  $B(S_{ij})$  then the coordinates

$$\vec{\mu}^{ij} = \begin{pmatrix} (u_x^{ij} - v_x)/w \\ (u_y^{ij} - v_y)/h \end{pmatrix} \quad (2)$$

$S$ 'nin ölçekten bağımsız bir temsilini sağlamaktadır.  $m \times n$ 'nin  $k$  adet noktasını örneklememiz bu sayede  $S$  şeklinin bir vektör olarak sunulmasını sağlamaktadır;

$$\vec{r} = [\mu_x^{i_1 j_1}, \mu_y^{i_1 j_1}, \dots, \mu_x^{i_k j_k}, \mu_y^{i_k j_k}] \in \mathbb{R}^{2k} \quad (3)$$

( $i_a < i_b$  eğer  $a < b$ ).

Dikkate alınması gereken bir nokta, yapılan bu gömme ( $\mathbb{R}^{2k}$ 'da) ölçekten bağımsız iken, dönüşlere karşı bağımsız değildir. Ancak, olağan dışı yürüyüş tespitinin gerçek dünya uygulamaları için insan silüetlerinin resim yüzeyi içinde yer alacak bir kamera sisteminin varsayılması makul olanıdır( **Ne demek istiyor?**) (Şekil-1'deki gibi). Bu sebeple, genelde yürüyen insanların şekillerinin **dik ( uprighth ? )** olması yüzünden dönüşlere olan bağımsızlık en önemli sorun teşkil eden mesele olmamaktadır. **Rather, exceptional feature vectors that result from mapping a lattice onto a somewhat rotated silhouette will be a strong indicator of abnormal gait.**

### 3. Yürüyüş sınıflandırması için özellikler(feature)

Bölüm 2'de verilen ihtiyaçlara karşılık verebilecek bir  $m \times n$  dizi boyutu seçmek keyfi(arbitrary) bir işlemdir. **Obviously, too fine a grid would be more sensitive against distorted shapes or individual traits than a coarser one.** Çok kabaca yapılmış bir kafeslendirme önemli silüet dinamiklerini yakalayamayacaktır. Yapılan deneylerde, 140 ila 240 arasında hücre içeren dizilerin makul doğrulukta sınıflandırmalar yaptığı görülmüştür[bauckhage05]. Bu gözlemlere

provide a scale invariant representation of  $S$ . Sampling  $k$  points of an  $m \times n$  lattice therefore allows to represent the shape  $S$  as a vector

$$\vec{r} = [\mu_x^{i_1 j_1}, \mu_y^{i_1 j_1}, \dots, \mu_x^{i_k j_k}, \mu_y^{i_k j_k}] \in \mathbb{R}^{2k} \quad (3)$$

where  $i_a < i_b$  if  $a < b$  and likewise for the index  $j$ . Note that while this embedding in  $\mathbb{R}^{2k}$  is scale invariant, it is not invariant against rotations. However, for real world applications of abnormal gait detection, it is reasonable to assume a camera setup where human silhouettes are within the image plane (as are the examples in Fig. 1). Therefore, since shapes of walking people will usually appear to be upright, rotation invariance is not of primary concern for our application. Rather, exceptional feature vectors that result from mapping a lattice onto a somewhat rotated silhouette will be a strong indicator of abnormal gait.

### 3. Features for gait classification

Choosing array dimensions  $m \cdot n$  that cope with the requirements listed in Section 2 seems an arbitrary task. Obviously, too fine a grid would be more sensitive against distorted shapes or individual traits than a coarser one. A meshing that is too coarse, however, might not capture essential silhouette dynamics. From experimenting with a standard database of shapes, we found that arrays of sizes between a 140 and 240 cells allow for robust classification [bauckhage05].

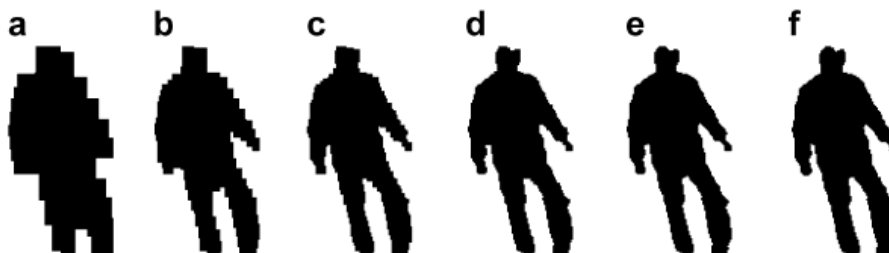


Fig. 3. Approximation of a binary shape by bounding box arrays of increasing dimensions  $m \times n$ . Given the number of columns  $n$ , the number of rows  $m$  in these examples was determined according to Eq. (1). (a)  $8 \times 4$ ; (b)  $16 \times 8$ ; (c)  $32 \times 16$ ; (d)  $64 \times 32$ ; (e)  $128 \times 64$ ; (f) original.

dayanarak 16 x 10'luk bir parselleme boyutu aşağıda anlatılan deneylerde kullanılmak üzere seçilmiştir. Seçilen bu değer güzel sanatlarda da kullanılan 3:5:8, kafa, gövde ve bacakların oranını yansıtan ölçümlerle desteklenmiştir. Şekil 6(b)'nin solunda gösterilen iki şekil, dik yürüyen iki insan için bu değerlerin uygun bir oranı temsil ettiğini izah etmektedir.  $m=16$  (3+5+8) seçiminde kullanılan bu estetiksel nedenden sonra, sütun sayısının belirlenmesinde de aynı yöntemin izlendiğini söylemek şaşırtıcı olmaz. 3:5 ve 5:8 yaklaşık olarak altın oranı vermektedir (1.618). Bu yüzden  $n=10$  seçilerek insan silüetleri için uygun bir kutu (container) olacak altın dikdörtgen -kenar oranı '1:altın oran' olan dikdörtgen- sağlanmıştır.

Bu kafes boyutlarını kararlaştırdıktan sonra, hangi noktaların örneklenerek yürüyüş sınıflandırılması için uygun vektör temsili sağlanabileceğimizi bulmamız gerekiyordu. Standart veri setleriyle yapılan deneyler, kafesin sınırlarından seçilen belirli noktaların doğru sınıflandırmayı mümkün kıldığını ortaya koymuştur [bauckhage05]. Dahası, çeşitli insan şekilleri üzerinde yapılan araştırmalarda, kafa omuz ve bacaklar arasındaki ilişkiler dik ve normal bir yürüyüş ile herhangi bir yalpalayan yürüyüş arasındaki farkları ortaya çıkaran önemli işaretler içermektedir. Şekil-6(b) buna bir örnek teşkil etmektedir. Soldaki iki şekil normal yürüyen iki kişiyi göstermektedir. İki durumda da kafa, omzun bir hayli üstünde, omuzlar ve kollar neredeyse bir doğru açıyı kapsar (180 dereceden büyük olması gerekiyor cümleyi tam anlamadım) ve ayaklar birbirine yakın durumdadır. Sağdaki iki şekil ise yalpalayan iki insanın video görüntülerinden çıkarılmış anları göstermektedir. Burada kafa ve omuz ile beraber bacakların birbiriyle ilişkisi tamamen rastgeledir. Dahası Şekil-1'den de görülebileceği üzere, normal yürüyüş kolların ve bacakların neredeyse aynı frekansta olan periyodik hareketlerini içerirken, yalpalayan yürüyüşteki hareketler aperiodyik ve rastgeledir. Bu bölgelere ait (kafa, omuz ve bacaklar), örneklenmiş kafes noktaları, yürüyüş sınıflandırmasında bize çok değerli bilgiler verecektir.

According to these empirical findings, a grid size of 16 · 10 was chosen for the experiments reported below. This particular choice was further motivated by experiences from the Fine Arts: the ratio 3:5:8 is often considered to be a pleasing measure of the relative sizes of head, torso and legs of the human body. The two shapes on the left of Fig. 6(b) illustrate that this rule of thumb indeed is reasonable for upright silhouettes. Given this aesthetic basis for choosing  $m = 16$  rows, the choice for the number of columns followed a similar path. Observe that 3:5 and 5:8 are approximations of the golden ratio / 1:1.618 . . . So we chose  $n = 10$  to approximate a golden rectangle - a rectangle with a side ratio of 1 : / and which is often considered a pleasing container of human silhouettes.

Having decided the lattice dimensions, it remains to determine which lattice points to sample in order to provide a suitable vector representation for gait classification. Experiments with a standard database of shapes revealed that in particular the points on the boundary of the lattice enable correct classification [bauckhage05]. Moreover, from inspecting exemplary series of human shapes, it appeared that relations between head and shoulders and between the feet provide reliable cues to distinguish normal upright walking from any kind of wobbling. Fig. 6(b) exemplifies this. The two shapes on the left correspond to instances of the gait cycle of two subjects walking normally. In each case, the head is well above the shoulders, shoulders and arms nearly enclose a right angle and the feet are not too far apart. The silhouettes on the right, in contrast, were extracted from video sequences showing abnormal gait. Here, the relation between head and shoulders as well as between the feet varies arbitrarily. Moreover, as can be seen in Fig. 1, normal gait is also characterised by periodic motion of almost constant frequency of arms and feet, whereas abnormal movements are aperiodic and random. A sample of lattice points that correspond to those parts of a silhouette which depict the head and shoulders and, therefore, the arms and feet will provide valuable information for gait classification.



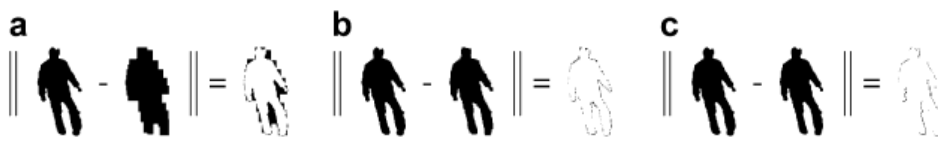


Fig. 4. Visualisation of shrinking Hamming distances  $d_H$  between a shape  $S$  and box array approximations  $B$  of growing array dimensions  $m \times n$ . (a)  $8 \times 4$ ; (b)  $32 \times 16$ ; (c)  $128 \times 64$ .

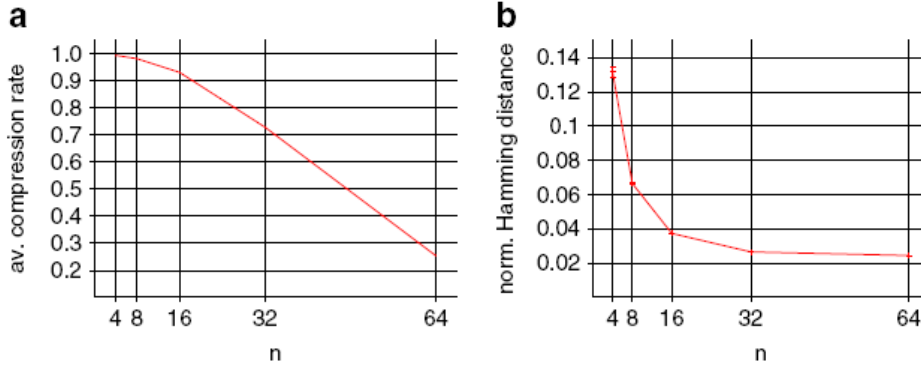


Fig. 5. Average compression rates and normalised Hamming distances obtained from shape encoding using box arrays of dimensions  $m \times n$ . The number of columns  $n$  was chosen as indicated; the corresponding number of rows  $m$  resulted from Eq. (1).

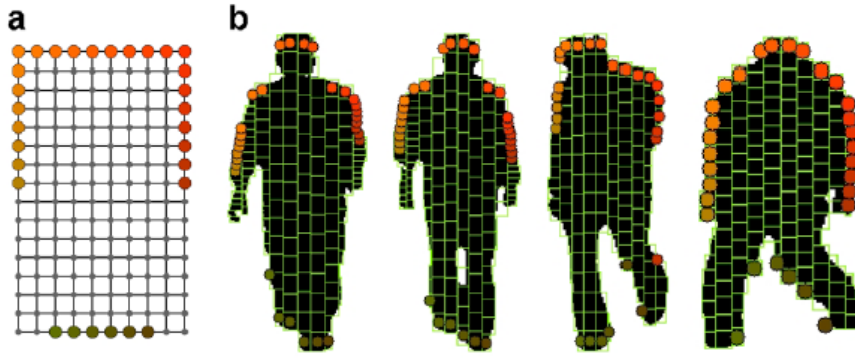


Fig. 6. A sample of  $k = 30$  points on a  $16 \times 10$  lattice and examples of how this sample is mapped onto different silhouettes.

Çalışmamızın bu noktası, 2. Bölümde anlatılan basit olarak yapılmış, şeklin 2D'lu kafesleme işleminin değerini tam olarak ortaya koymaktadır. Bu haritalama işleminin topolojisi gereği, kafesin üst ve alt sınırlarından (ayrıca sağ sol sınırları) örneklenen noktalar genellikle kafa, omuz, kollar ve bacaklara denk gelmektedir. Şekil-6, bizim kullandığımız 30 kafes noktasını ve onların farklı örnekler üzerinde denk geldikleri noktaları göstermektedir. Buradan da açıkça görülmektedir ki, normal dik bir yürüyüş için bu noktalar, uzuvların takip edilmesine olanak sağlamaktadır. Ne takip algoritmaları, ne de tanıma işlemleri yapılmadan bu belirtilen vücut parçalarının tespiti yapılabilir.

Yürüyüş zamana bağlı(temporal) bir olay olduğu için tek bir örnekten alınan bilgiler yürüyüşün normal ve ya anormal olduğu hakkında bilgi vermeye yeterli olmayacaktır.

At this stage of our discussion, the simple mapping between 2D lattices and shapes introduced in Section 2 reveals its full potential. Due to the mapping's topology preserving properties, a sample of points with lattice coordinates on the upper and lower border of the lattice will usually, i.e., in the case of upright silhouettes, correspond to head, shoulders, arms, and feet. Fig. 6 displays the  $k = 30$  lattice points we used in our experiments and where they are located on different silhouettes. Obviously, for normal gait, this scheme allows to roughly keep track of limbs. Neither feature tracking algorithms nor recognition procedures have to be applied to infer the location of significant body parts. Since gait is a temporal phenomenon, information about a single instance of a gait cycle might not be sufficient to determine whether normal or abnormal gait is being observed.

Bu zamana bağılı olayları sınıflandırma işlemine dahil etmek için her  $t$  zaman basamağında sıralanmış özellik vektörlerini her  $t(1) \in \{t, \dots, -\Delta\}$  olacak şekilde,

$$\vec{s}_t = \vec{r}_t \oplus \vec{r}_{t-1} \oplus \dots \oplus \vec{r}_{t-\Delta} \quad (4)$$

ve,

$$\vec{r}_t = [\mu_x^{k,k}(t_t), \mu_y^{k,k}(t_t)], \quad k = 1, \dots, 30 \quad (5)$$

dir.

Bu yüzden şekil ve onun yakın geçmişi yüksek boyutlu bir özellik vektörünün ortalamaları şeklinde karakterize edilir;  $\vec{s}_t \in \mathbb{R}^{2k(\Delta+1)}$ . Bir sonraki bölümde açıklanacak olan deneylerde  $\Delta$  sırasıyla, 0, 10 ve 20 olarak seçilmiştir. Örnek olarak o anki ve ondan sonra gelecek olan kesitlerin (frame) 60 boyutlu şekil tanımlayıcıları 1260'a kadar boyutları çıkabilen  $\vec{s}_t$  özellik vektörlerine birleştirilmiştir. (??)

#### 4. Deneyler

Yaptığımız yaklaşımın olağan dışı yürüyüş tespitinde uygulanabilirliğini test etmek için bir grup yeşil ekran önünde kaydedilmiş videoyu kullandık. Yedi kişi önce normal daha sonra da alışlagelmişin dışında (örn: baş dönmesi gibi bir denge sorunu yaşadıkları durumu varsayarak) yeşil ekran önünde yürütüldü. Elimizde olan kamera düzeneği ve yürüme platformuna bağlı olarak, kişiler ya kameraya doğru ya da kameradan uzaklaşacak şekilde, kamerayla  $0^\circ$  ila  $45^\circ$  arasında bir açı yaparak yürüdüler.

Yürüyen insanların olmadığı 50 kesitten basit bir istatistiksel model (statistical model) kullanılarak arka planın renk dağılımını hesaplayan hareket segmentasyonu yapıldı. Daha sonra  $5 \times 5$  'lik bir median filtreme işlemini takip eden birleşik birleşen analizi (connected components analysis), bu makalede gördüğümüz şekillerin elde edilmesini sağladı.

In order to incorporate temporal context into the classification process, at each time step  $t$ , we consider concatenated feature vectors

$$\vec{s}_t = \vec{r}_t \oplus \vec{r}_{t-1} \oplus \dots \oplus \vec{r}_{t-\Delta} \quad (4)$$

where, for each  $t(1)$ , we have  $t(1) \in \{t, \dots, -\Delta\}$

Therefore, a shape and its recent history will be characterised by means of a high dimensional feature vector  $\vec{s}_t \in \mathbb{R}^{2k(\Delta+1)}$ . For the experiments described in the next section,  $\Delta$  was set to 0, 10 and 20, respectively, i.e., the 60 dimensional shape descriptors of the current frame and its predecessors were combined into feature vectors  $\vec{s}_t$  of up to 1260 dimensions.

#### 4. Experiments

In order to test the feasibility of our approach to abnormal gait detection, we experimented with a set of videos recorded in front of a green screen. Seven subjects were asked to walk in a normal fashion as well as in an unusual way, e.g., as if they were suffering a balance deficiency such as dizziness. In accordance with the setting of our application domain, the movement in these baseline sequences is either towards or away from the camera, and the angle to the camera varies between  $0^\circ$  and about  $45^\circ$ .

Motion segmentation was done using a simple statistical model where 50 frames of video without walking people were used to determine the colour distribution of the scene background. Subsequent filtering using a  $5 \times 5$  median followed by a connected component analysis provides shapes, such as those seen throughout this paper.

Amacımız olan, normal ve olağan dışı yürüyüşleri sınıflandırılması için bir çok örüntü tanıma tekniği uygulanabilir. Biz radyal temelli bir kernel fonksiyonlu (bu nedir??) destek vektör makinesini (DVM) kullanmayı uygun bulduk. Bunu kullanmamızdaki sebep iki sınıflı problemlerde yüksek güvenilirlikte sonuçlar verebilmesiydi [vapnik05]. Ayrıca, bu sınıflandırıcıların destek vektör özellikleri tek tarafa yoğunlaşmış eğitim setlerinde de iyi sonuçlar vermektedir. Gerçek dünyada normal yürüyüşlerin olağan dışı yürüyüşlerden fazla olacağı düşünüldüğünde bunun da önemli bir etken olduğu anlaşılmaktadır. Bazı sınıflandırıcılar bu nedenle daha fazla örnek içeren sınıfı işaret eden yanlış sonuçlar verirken, sınıf sınırlarını karakterize eden eğitilmiş bir DVM, kolayca eğitim örneklerine bağlı olarak tahminler yapmaktadır. Ancak, bir DVM'sini eğitmek, ikinci dereceden bir optimizasyon problemini çözmeyi gerektirmektedir. Bu işlem elimizde yüksek boyutlarda bir özellik uzayı ve çok sayıda eğitim örneği içeren bir veri seti olduğunda karmaşık bir probleme dönüşmekteydi. Bunu göz önünde bulundurduğumuz için DVM<sub>light</sub> algoritmasını uyguladık [joachims99]. Bu, eğitim sürecini küçük seri işlemlere ayırmaktaydı(??).

Tablo-1(a) üç seriden oluşan **baseline** deneylerin eğitim parametrelerini özetliyor. (??) Her seride 4 kişiyi gösteren, 7 video eğitim için kullanılmıştır. Bu videolardan dördü normal

Concerning our goal of distinguishing between normal and abnormal gait, many techniques developed in the field

of pattern recognition are applicable. We opted for support vector machines (SVMs) with a radial basis kernel function because they are known to perform very reliably in twoclass problems [vapnik05]. Also, the support vector property of these classifiers copes with possibly biased training data: in real world applications of gait analysis, examples of normal gait will outnumber those of abnormal gait. While some classifiers thus may favour the class with more examples, a trained SVM simply predicts according to those training examples which characterise the class boundaries. However, training an SVM requires solving a quadratic optimisation problem, which can be burdensome if the considered feature space is of high dimensionality and there are many training examples. We therefore applied the SVMlight algorithm [joachims99], which tackles these issues by decomposing the training process into a series of smaller tasks.

Table 1  
Overview of parameters and results for training and testing abnormal gait detection

$\Delta$	No of frames normal	No of frames abnormal	Leave one out error (%)
(a) Training Parameters and results			
0	1359	1128	26.5
10	1323	1101	11.6
20	1283	1071	6.1
$\Delta$	Gait	No of frames	Accuracy (%)
(b) Test Parameters and results			
0	Normal	1227	73
	Abnormal	1413	61
10	Normal	1195	77
	Abnormal	1382	70
20	Normal	1157	72
	Abnormal	1350	82

yürüyüş, üçü ise olağan dışı yürüyüş örneklerini içermektedir. Her yürüyüşün kendi içinde eşsiz olması sebebiyle videolar farklı sayıda kesitten oluşmaktaydı (sonuçta herkes bir mesefayı aynı sürede yürüyemiyor. Hatta aynı insan bile farklı sürede aynı mesefayı yürütebilir. ) Zamana bağlı bir şart olarak (Choosing a temporal context),  $\Delta=20$  seçilerek, her eğitim örneğinden ilk yirmi kesiti çıkartıldı. Tablonun son sütunundaki figureler(?) DVM'sinin tahmini başarısının sonuçlarını vermektedir. (??) 'Bir tanesini dışarıda bırak' (leave one out) testini her bir bağımsız kesite uyguladığımızda listelenen hata oranlarını bize geri döndürdü. Hata ortadan kaybolmadığı için, normal ve olağan dışı yürüyüşlerin özellik uzay alanlarının üst üste bindiği aşıkardı. Ancak, eğer zamana bağlı şartlar artırılırsa, örneğin, eğer özellik uzayının boyutları büyürse, bu üst üste binme azalır ve buna bağlı olarak hata oranı düşer. Bir insanın normal ve ya olağan dışı yürüdüğüne algılanması için daha uzun bir gözlemin, daha yüksek doğrulukta bir sonuç vereceğini sezgisel olarak algılayabiliriz. Ayrıca daha yüksek boyutta bir vektör uzayı, daha dağınık bir dağılım sağlayacak ve sınıflandırma için daha geniş bir alan olacaktır. (??) Zaman bağlı şartları arttırdığımızda hata oranı daha da azalacaktır. Ancak uygulamalarda oluşacak kısıtlamalar sebebiyle, bu seçeneği düşünmedik. Bunu nedeni; ilk olarak 'boyutların laneti' (curse of dimensionality) sebebiyle doğrusal olarak artan özellik vektörlerinin üstel olarak artan eğitim verileri gerektirmesiydi. İkinci olarak ise, yüksek boyutlardaki özellik vektörlerinin DVM'si tabanlı sınıflandırmalarda hatırı sayılır bir şekilde işlem emeğini arttırmasıydı (örn: [kienze05]). Bu da gerçekzamanlı uygulamalar için kullanılabilirliği

Table 1(a) summarises the training parameters of three series of baseline experiments. In each series, the same set of 7 videos showing 4 individuals was used for training.

Four of these videos show normal gait. The remaining 3 contain examples of abnormal gait. Due to the varying length of temporal context, the number of frames available for training differs in each experiment. Choosing a temporal context of  $D = 20$ , for instance, implies that the first twenty frames of each training sequence have to be skipped from training. The figures in the last column of the table are an estimate of the quality of the resulting SVM. Performing a 'leave one out' test on the individual frames of the training data produced the listed error rates. Since the error rates do not vanish, the feature space areas of normal and abnormal gait obviously overlap. However, if the temporal context increases, i.e., if the dimension of the feature space grows, the overlap decreases and thus the error rate decreases. This agrees with intuition and mathematical wisdom: the decision of whether or not a person moves abnormally is more reliable the longer the person is being observed. Also, in higher dimensional vector spaces, data will be more sparse, and there will be more space for partitioning. Further increasing the temporal context will further reduce the error rates. However, due to practical constraints imposed by our application scenario, we did not consider this option: first of all, due to the 'curse of dimensionality', linearly growing feature vectors require an exponentially growing amount of training data. Second of all, very high dimensional feature vectors considerably increase the computation effort of SVM-based classification (cf. e.g., [kienze05]) and thus threaten the real time capability crucial for our system.

sekteye uğratmaktaydı.

Bağımsız test setinden elde edilen sonuçlar ve parametreler Tablo-1(b)'de gösterilmektedir. Burada, 3'ü eğitim setine dahil edilmemiş 5 bireyin 7 videosunu ele aldık. Bu 7 test videosundan 4'ü olağan dışı yürüyüşe, 3'ü ise normal yürüyüşe sahip insanları göstermekteydi. Farklı zamana bağlı şartlardan dolayı (due to the different temporal context), kesit sayısı deneyden deneye değişmekteydi. Dikkati çeken bir olay ise, artan zamana bağlı şartlar olağan dışı yürüyüş tespitini geliştirirken, normal yürüyüş tespitine etkisi tutarsız oldu. Genel olarak, özellikle  $\Delta=20$  seçildiğinde, kesit başına algılama başarısı kabul edilebilir düzeyde olmuştur. Ancak, yanlış pozitif oranı olağan dışı yürüyüş tespitinin gerçek dünya uygulamalarında kullanılabilmesi için olması gerekenden yüksek olmaktadır.

Bu problem için bizim çözümümüz Şekil-7 ve 8'de tarif edilmiştir. Karışık, muhtemelen sabit olmayan arka plan perdesi (scene doğru?) ve homojen olmayan parlaklık dağılımlarıyla başa çıkmak için, hareket segmentasyonu işlemi, kernel yoğunluk tahmini (estimators)? yöntemi, arkada plan modellemesi için kullanılmıştır [elgammal00]. Bu sayede daha iyi bir arka plan modeli oluşturulmuştur. İki şekilde, öncelikle olağan dışı daha sonra normal olarak yürütülen 2 kişinin yaklaşık 30 kesitlik bir periyodunu kapsamaktadır. Şekillerin sağındaki renklendirme son 30 kesit içinde 'yalpalayan yürüyüş' olarak sınıflandırılmışların yüzdesini görselleştirmektedir. Şüpheli kesit sayısı arttıkça, kırmızı renklendirme artmaktadır

Parameters and results obtained on an independent test set are shown in Table 1(b). Here, we considered 7 videos of 5 individuals, 3 of which were not present in the training set. Out of the 7 test videos, 4 show instances of abnormal gait and 3 display usual walking behaviour. Again, due to the different temporal context, the number of frames varies among the experiments. What is noticeable is that, while a growing temporal context considerably improves the detection of abnormal gait, its effect on normal gait detection is inconsistent. Overall, especially for a temporal depth of

$D = 20$ , the per-frame recognition accuracy can be considered acceptable. However, the rate of false positives seems to be too high to be used in real world applications of abnormal gait detection. Our solution for this problem is depicted in Figs. 7 and 8. Note that, in order to deal with cluttered, possibly nonstatic scenes and inhomogeneous illumination conditions as in these examples, the motion segmentation procedure from our baseline experiments was improved using kernel density estimators for background modelling as proposed in [elgammal00]. Both figures cover a period of about 30 frames, taken from a sequence of a subject walking abnormally and normally, respectively. The idea of colouring the segmented shapes on the right of each panel was developed in collaboration with a partner from industry. It indicates the percentage of frames that were classified abnormal during the last 30 frames. The higher the number of dubious frames rises, the higher is the level of the red colouring. At the beginning of the part of the sequence that is shown in Fig. 7, only a single frame has been classified as abnormal gait.



Fig. 7. Example of abnormal gait detection over a period of 32 frames; the temporal context for classification was set to  $\Delta = 20$  frames. The colouring of the segmented shapes indicates the percentage of silhouettes in the past 30 frames that were classified to depict *abnormal* gait. Thus at the beginning of this sequence, the subject's gait appeared to be fairly normal whereas at the end each of his last 30 instances has been classified as abnormal gait.



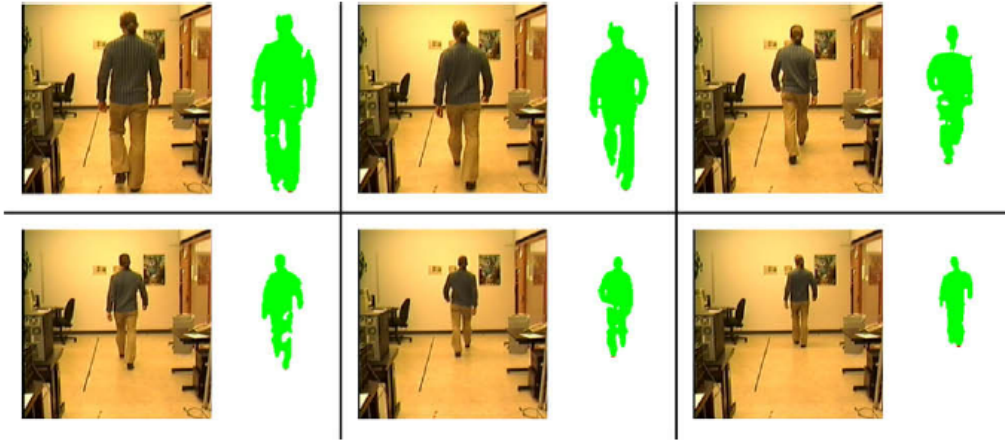


Fig. 8. Example of gait analysis over a period of 31 frames; again, the temporal context for classification was set to  $\Delta = 20$  frames. The sequence shown here is an example of normal gait. None of the 31 frames covered in this figure was recognised as an instance of abnormal gait. Correspondingly, the segmented shapes on the right of each panel are homogeneously coloured in green

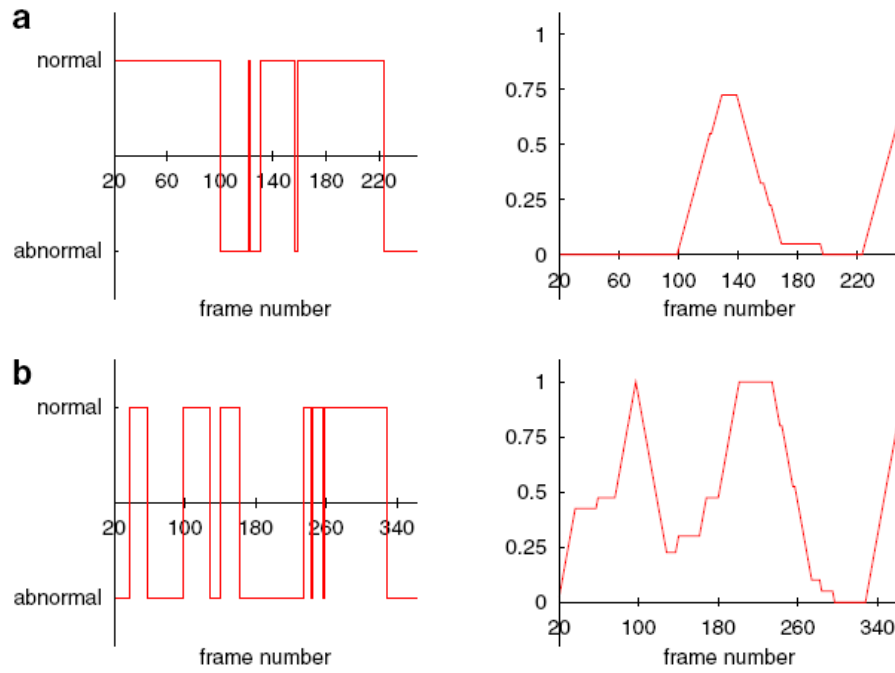


Fig. 9. Per frame classification results and corresponding percentage of dubious frames within the last 30 frames for two short sequences of normal and abnormal gait. (a) Example of a sequence of normal gait; (b) example of a sequence of abnormal gait.

. Şekil-7'deki yürüyüşün başında sadece tek bir kesit yalpalayan yürüyüş olarak sınıflandırılmıştır. Kişinin yalpalayan hareketlerinin devam etmesi neticesinde, son kesite gelindiğinde bu oran %100'e ulaşmıştır. Şekil-7'deki sağ alt panele baktığımızda bunu görebilmekteyiz.

Bu bağlamda, renklendirme zamana bağlı içerik(temporal context)'in daha yüksek bir

Corresponding to the wavering movement of the subject, however, the percentage of abnormal frames rises continuously throughout the sequence and reaches 100% in the panel at the bottom right. In this manner, the colour coding illustrates the use of temporal context on a higher level of abstraction.

seviyede kullanılmasını izah etmektedir. Bunu kesit tabanlı sınıflandırma sonuçlarına uygulanan geçici bir filtre olarak görebiliriz. Bir ve ya birkaç kesit yanlış sınıflandırılırsa bile, genel kesit sınıflandırmasının üzerindeki etkisi düşük olmaktadır. Bu olay Şekil-9'da örneklenmiştir. Tablo-1(b)'de, iki kişinin kesitlerinin teker teker sınıflandırılması ve zamana bağlı filtrelene işleminin ( context temporal) sonuçları gösterilmiştir. İki örnek içinde sınıflandırmada kullanılan zamana bağlı şart  $\Delta = 20$  kesit , ve takip eden geçici pürüzsüzleştirme(subsequent temporal smoothing) içinse, 30 kesit seçilmiştir. İki durumda da dikkate alınacak sayıda hatalı sınıflandırılmış kesitler bulunmaktadır ancak bunların devamlılığı bazen tek bir kesit olacak kadar kısa olmuştur. (üst üste yanlış sınıflanan frame'ler olmamıştır ki bu sayede daha önce bahsedilen kırmızı şekil hep yüksek dolulukta kalmıştır) Bu nedenle, son 30 kesite bakıldığında normal yürüyüşü gösteren dizide olağan dışı olarak şüphelenilen kesitlerin oranı %75'i hiçbir zaman geçmemiştir (%75'te az bir oran mı tartışılabilir) . Anormal yürüyüşü gösteren dizide ise, bazı bölümlerin şüpheli kesit yüzdesi bu sınırın üstüne bariz bir şekilde çıkmıştır. Gerçek dünya gözetim uygulamalarında, verilen bir zaman aralığında belirli bir eşik aşıldığında olağan dışı yürüyüş alarmı verilebilir.

## 5. Özet ve Genel Görüş

Bu makale otomatik yürüyüş analizini yürüyüş örüntülerine bakarak normal mi değil mi sorusuna cevap bulmak için ele aldı. Görsel yürüyüş analizine yapılan diğer katkılardan farklı olarak, bu makalede üstesinde gelinen problem, with in this paper requires a representation that abstracts from individual gait characteristics but allows for the classification of gait across individuals. Bu ihtiyacın doğrultusunda, 2 boyutlu kafeslerle insan silüeti arasında bir benzerşekillilik yaratıldı. Bu, silüetlerin doğru bir şekilde vektör uzayına gömülmesini sağladı. Uygun olan kafes noktaları örneklenerek uzuvların hareketleri hiçbir uzuv tanıma yöntemi uygulanmadan takip edilebildi.

Combining shape representations derived from several frames into lager feature vectors

It can be seen as a temporal filter that acts on the results of framewise classification. If one or several frames of a sequence are misclassified, it will have little effect on the general tendency or confidence level that becomes apparent from temporal filtering. This is further illustrated in Fig. 9 which depicts frame-wise classification results and the corresponding results of the temporal context filter for two of the test sequences considered in Table 1(b). In both cases, the temporal context for classification was set to  $D = 20$  frames and 30 frames were considered for subsequent temporal smoothing. For both sequences, there are considerably many misclassified frames, which, however, only occur over short periods of time, sometimes as short as as ingle frame. Consequently, for the sequence showing normal gait, the percentage of dubious frames within the last 30 frames never exceeds about 75%. For the sequence showing abnormal gait, on the other hand, there are several intervals where the percentage of dubious frames clearly exceeds this mark. For real world surveillance applications, it is thus possible to trigger an alarm once a certain level of abnormal gait has been detected over a given period of time.

## 5. Summary and outlook

This paper considered automatic gait analysis as a means to deduce if an observed walking pattern appears to be normal or not. In contrast to most contributions to visual gait analysis, the problem dealt with in this paper requires a representation that abstracts from individual gait characteristics but allows for the classification of gait across individuals. Addressing this requirement, we presented a homeomorphism between 2D lattices and shapes that enables a robust vector space embedding of silhouettes. Sampling suitable lattice points allows to roughly track the movement of limbs without requiring any limb recognition strategy.

Combining shape representations derived from several frames into lager feature vectors provides temporal context for the classification task. Experimental results underline that gait classification using support vector machines yields satisfiable accuracy. Temporal filtering of the classification results in further

provides temporal context for the classification task.

Yapılan deneyler sınıflandırmada destek vektör makinesi kullanıldığında tatmin edici sonuçlar alındığını göstermektedir. Geçici filtreleme ile (temporal filtering ??) ara sıra ortaya çıkan hataların etkisi ortadan kaldırılmaktadır.

Şu an da, yürüyüş sınıflandırılmasının karmaşık olmayan gerçek dünya düzeğinde güvenilirliğini ve uygulanabilirliğini, otomatik olarak arkada plan modelini çıkaran bir hareket segmentasyon tekniği kullanarak, göstermeye çalışmaktayız. Ek olarak, sistemimizin birden çok insanın birbirinin önüne geçeceği durumlar için adapte olabilmelerini sağlamaya çalışıyoruz. Bu amaçla, [okuma04]'te yer alan güçlü bir takip tekniğini denemekteyiz. Bu teknik bizim birkaç silüeti aynı anda çıkartmaya yarayacaktır.

improvements of the reliability of the presented framework, because it lessens the effect of sporadic misclassifications.

Currently, we have been working to demonstrate the applicability and reliability of our gait classification scheme in different, unconstrained real world settings, using a motion segmentation technique that automatically acquires background models for non-rigid scenes. In addition, we are working to adopt our approach to situations with several people in such a way that they will occlude each other. To this end, we are experimenting with robust tracking techniques, as introduced in [okuma04], that allow us to extract and analyse several silhouettes simultaneously.